

Silvia Galvão de Souza Cervantes

**Um Algoritmo Descentralizado
para Controle de Tráfego Urbano em Tempo Real**

**FLORIANÓPOLIS
2005**

UNIVERSIDADE FEDERAL DE SANTA CATARINA
CURSO DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

**Um Algoritmo Descentralizado para Controle
de Tráfego Urbano em Tempo Real**

Dissertação submetida à
Universidade Federal de Santa Catarina
como parte dos requisitos para a
obtenção do grau de Doutor em Engenharia Elétrica.

Silvia Galvão de Souza Cervantes

Florianópolis, Março de 2005.

Um Algoritmo Descentralizado para Controle de Tráfego Urbano em Tempo Real

Silvia Galvão de Souza Cervantes

‘Esta Tese foi julgada adequada para a obtenção do título de Doutor em Engenharia Elétrica, Área de Concentração em *Controle, Automação e Informática Industrial*, e aprovada em sua forma final pelo Programa de Pós-Graduação em Engenharia Elétrica da Universidade Federal de Santa Catarina.’

Werner Kraus Junior
Orientador
Presidente

Eduardo Camponogara
Co-orientador

Prof. Alexandre Trofino Neto
Coordenador do Programa de Pós-Graduação em Engenharia Elétrica

Banca Examinadora:

Dr. Ademar Ferreira

Dra. Maria Alice Prudêncio Jacques

Dr. José Eduardo Ribeiro Cury

A Deus, o único que está acima da nossa ignorância e não me deixou sozinha em mais esta caminhada.

AGRADECIMENTOS

Ao meu marido José Augusto pelo apoio e paciência e ao meu pequeno filho Pedro Augusto que não entendia o porquê de tantas recusas aos pedidos de atenção.

Ao meu orientador Werner pela paciência, compreensão, dedicação e apoio. Ao meu co-orientador Eduardo pela ajuda e dedicação.

A minha mãe pelo carinho. Minhas irmãs Célia, Márcia e Sônia que mesmo distantes não me deixaram sem palavras de apoio e incentivo.

A todos, que não são poucos, que direta ou indiretamente colaboraram para que este trabalho se tornasse uma realidade.

Resumo da Tese apresentada à UFSC como parte dos requisitos necessários para obtenção do grau de Doutor em Engenharia Elétrica.

Um Algoritmo Descentralizado para Controle de Tráfego Urbano em Tempo Real

Silvia Galvão de Souza Cervantes

Agosto/2005

Orientador: Werner Kraus Junior

Co-orientador: Eduardo Camponogara

Área de Concentração: Controle, Automação e Informática Industrial

Palavras-chave: Controle Semafórico em Tempo Real, Tráfego Urbano, Otimização

Número de Páginas: 108

Este estudo busca contribuir para a melhoria da circulação de tráfego em redes viárias através do controle ótimo dos semáforos. O critério a ser minimizado é o atraso veicular médio a que estão submetidos os motoristas. Foi desenvolvido um modelo de tráfego baseado em equações dinâmicas que descrevem o comportamento do sistema viário a partir de contagens veiculares. As equações consideram características do acoplamento entre interseções e permitem a análise dos efeitos de diferentes políticas de controle.

Com base no modelo, foi desenvolvido um algoritmo de busca em profundidade para realizar o controle em tempo real, em uma configuração de controle preditivo descentralizado com horizonte deslizante. O método de controle fornece, a partir de contagens de fluxo veicular, os tempos de abertura dos semáforos que resultam no melhor desempenho possível para a malha viária. A atualização da ação de controle é feita a cada 4 s para compensar, via realimentação da informação dos detectores veiculares, as possíveis imprecisões na modelagem e a natureza descentralizada do algoritmo.

Para fins de avaliação da qualidade da solução do método proposto, foi feita a transcrição do modelo para o formalismo de programação matemática, como um programa linear inteiro misto. A formulação permite o uso de pacotes de otimização para obtenção de soluções ótimas ou próximas do ótimo, servindo como referência para as soluções com a heurística do controle preditivo. Também, foram usados resultados de tempos fixos do programa Transyt para comparações com ambos os métodos anteriores.

Resultados de simulação indicam a viabilidade do método proposto, em termos da qualidade da solução frente à solução ótima global, e sua superioridade em relação a planos de tempo fixo.

Abstract of Thesis presented to UFSC as a partial fulfillment of the requirements for the degree of Doctor) in Electrical Engineering.

A Decentralized Algorithm for Urban Traffic Control in Real Time

Silvia Galvão de Souza Cervantes

August, 2005

Name of the advisor: Werner Kraus Junior

Name of the co-advisor: Eduardo Camponogara

Area of Concentration: Control, Automation and Industrial Computing

Key words: Real Time Signal Control, Urban Traffic, Optimization

Number of Pages: 108

A contribution for the improvement of vehicular traffic flow in road networks via the optimal control of traffic lights is presented in this work. The minimization criterion is the average delay experienced by drivers. A traffic model is developed based on dynamic equations describing the road network behaviour from traffic count data. The equations take into account the coupling among intersections and allow for the analysis of control policies.

Based on the model, a depth-first searching algorithm is presented. It implements a real-time control policy in a decentralized, predictive control setting with rolling horizon. The control method calculates, from traffic data, timings for the traffic lights that result in the best possible performance for the road network. Control update is done every 4 s to compensate, via feeding back the information from vehicle detectors, for modelling errors and the decentralized nature of the algorithm.

To evaluate the quality of the solution from the proposed method, the model is translated as a mixed-integer linear program. The formulation allows the use of optimization packages that provide global optimal (or near optimal) solutions that serve as a reference for the assessment of the quality of the heuristic solution. Moreover, fixed-time setting obtained from the Transyt software are used for comparisons with both methods.

Simulation results indicate the viability of the proposed method in terms of the quality of the solution vis-a-vis the global optimal solution and its superiority when compared to fixed-time plans.

Sumário

1	Introdução	1
1.1	Definição do Problema	2
1.2	Objetivo da Pesquisa	2
1.3	Contribuição da Pesquisa	3
1.4	Organização do Documento	3
2	Conceitos Básicos de Tráfego Veicular Urbano	4
2.1	Introdução	4
2.2	Definições Preliminares	4
2.2.1	Redes Viárias	4
2.2.2	Fluxos Veiculares	5
2.2.3	Elementos de Controle de tráfego	7
2.3	Modelos de Tráfego	8
2.4	Tipos de Controle Semafórico	9
2.4.1	Tempo Fixo	9
2.4.2	Controle Atuado	10
2.5	Otimização	10
2.6	Coordenação entre interseções	12
2.7	Formação e Descarga de Filas	12
2.7.1	Modelo de Fila Vertical	14
2.7.2	Modelo de Fila Horizontal	16

2.7.3	Comparação entre os modelos de fila	18
2.8	Critério de Desempenho	19
2.8.1	Atraso	19
2.8.2	Número de Paradas	21
2.8.3	Comparação entre os critérios atraso e paradas	21
2.9	Conclusões	23
3	Revisão Bibliográfica	24
3.1	Introdução	24
3.2	Modelos de Tráfego e Estratégias de Controle Associadas	24
3.2.1	TRANSYT	29
3.2.2	TRANSYT-AUT	33
3.2.3	MAXBAND	33
3.2.4	SCOOT	34
3.2.5	SCATS	35
3.2.6	PRODYN	37
3.2.7	OPAC	39
3.2.8	RHODES	42
3.2.9	ALLONS-D	43
3.2.10	CRONOS	45
3.2.11	TUC	46
3.3	Classificação	48
3.3.1	Arquitetura e Estratégia de Otimização	49
3.3.2	Modelo de Fila e Detecção	49
3.3.3	Critério de Desempenho	51
3.4	Conclusões	51

4	Modelo de Simulação de Tráfego e Metodologia de Solução	53
4.1	Introdução	53
4.1.1	Modelo de Tráfego Desenvolvido para Este Trabalho	53
4.2	Algoritmo de Controle em Tempo Real	56
4.2.1	Horizonte Deslizante	57
4.2.2	Predição de Chegadas	58
4.2.3	Busca em Profundidade	58
4.2.4	Crítério de Desempenho	60
4.2.5	Coordenação Implícita	61
4.3	Resultados de Simulações	61
4.3.1	Características da Rede de testes	61
4.3.2	Comparações com o TRANSYT	62
4.3.3	Desempenho do Algoritmo de Controle em Tempo Real	65
4.4	Conclusões	68
5	Modelo de Tráfego em Programação Matemática	69
5.1	Introdução	69
5.2	Modelo para Programação Matemática	70
5.3	Resultados	73
5.4	Extensões ao Modelo de Programação Matemática	75
5.4.1	Restrição de Verde Máximo	75
5.4.2	Restrição de Vermelho Máximo	76
5.4.3	Coordenação entre semáforos adjacentes	76
5.4.4	Aproximação de Filas Horizontais	77
5.4.5	Discussão	79
5.5	Conclusões	79

6	Resultados para malha viária aumentada	81
6.1	Introdução	81
6.2	Características da Rede de Testes	81
6.3	Comparações entre Algoritmo Descentralizado e TRANSYT	82
6.4	Comparação do Algoritmo descentralizado com o Modelo Global	84
6.5	Conclusões	85
7	Conclusões e Perspectivas	89
7.1	Introdução	89
7.2	Conclusões	89
7.3	Perspectivas	90
8	Publicações	91
A	Modelo Matemático	92
A.1	Algoritmo de Programação Matemática	92
A.1.1	Arquivo de dados - 3 interseções	95
A.1.2	Arquivo de dados - 6 interseções	97
B	Distribuições	102
B.1	Distribuição Escada	102
B.2	Distribuição Exponencial	103
B.3	Distribuição Constante	103
B.4	Distribuição Pulsada	103
C	Configuração do sistema TRANSYT	105

Lista de Figuras

2.1	Interseção isolada.	5
2.2	Exemplo de um ciclo com sua divisão entre estágios.	8
2.3	Processo de formulação e validação de um modelo matemático.	11
2.4	Coordenação entre interseções de uma via.	13
2.5	Formação e Descarga de Filas.	14
2.6	Levantamento de trajetória e velocidade com aceleração e desaceleração instantânea .	15
2.7	Chegadas e Partidas de Veículos Acumuladas para um caso ideal (adaptada de [37]).	16
2.8	Modelo de fila vertical.	17
2.9	Fila horizontal (adaptada de [56]).	17
2.10	Comparação entre modelo de fila horizontal e vertical.	18
2.11	Componentes da formação do atraso (adaptado de [58].)	19
2.12	Componentes da formação do atraso sobre-saturado.	20
2.13	Exemplo de aplicação do critério de atraso.	22
2.14	Exemplo de aplicação do critério de parada.	22
3.1	Atraso médio veicular de uma faixa	26
3.2	Exemplo para o cálculo do <i>split</i> [12]	27
3.3	Diagrama espaço-tempo	29
3.4	Modelo TRANSYT	30
3.5	Exemplo de um Perfil de Fluxo Cíclico	31
3.6	Exemplo de um PFC, <i>IN-profile</i>	32
3.7	Exemplo de um PFC, <i>GO-profile</i>	32

3.8	Exemplo de um PFC, <i>OUT-profile</i>	32
3.9	Sistema utilizado pelo SCATS.	36
3.10	Cálculo do controle - Horizonte Deslizante (Adaptada de [22]).	38
3.11	PRODYN - Coordenação implícita	40
3.12	OPAC - Horizonte Deslizante (adaptado de [29])	41
3.13	Níveis de Coordenação - OPAC	42
3.14	RHODES - Arquitetura hierárquica	43
3.15	Via de uma rede urbana	46
3.16	Localização de detectores	50
4.1	Seções de uma faixa.	55
4.2	Árvore de busca - semáforo de dois estágios.	57
4.3	Rede de Tráfego para 3 interseções.	62
4.4	Comparação TRANSYT x algoritmo de descentralizado.	64
4.5	Comparação entre diferentes distribuições de chegadas.	66
4.6	Comparação entre diferentes distribuições de fluxo para chegadas pulsadas.	67
5.1	Grafo da rede de tráfego com 3 interseções.	71
5.2	Comparação algoritmo descentralizado (a) \times modelo global (b), carregamento desbalanceado médio, chegada em escada.	75
5.3	Rede para coordenação.	76
5.4	Valores hipotéticos de α_j	78
6.1	Rede de Tráfego.	82
6.2	Comparação algoritmo descentralizado (a) \times TRANSYT (b)	84
6.3	(a) Carregamento desbalanceado, alto, escada, (b) Carregamento balanceado, médio, escada.	87
6.4	Comparação algoritmo descentralizado (a) \times (b) modelo global.	88
B.1	Distribuição de chegadas para o primeiro arco arterial da rede.	103

Lista de Tabelas

3.1	Classificação dos Modelos de Tráfego e Estratégias de Controle.	51
4.1	Proporção de fluxos nas faixas.	62
4.2	Comparação entre valores de atraso com TRANSYT e estratégia descentralizada (pcu-h/h e pcu-s/s).	63
4.3	Valores de atraso (s) obtidos com algoritmo descentralizado.	65
5.1	Comparação do atraso total (s) entre algoritmo descentralizado (D) e modelo global (G). Simulação realizada para $25t_s$	74
5.2	Proporção de melhoria de desempenho do modelo global \times algoritmo descentralizado.	74
6.1	Atraso em $pcu - h/h$ (TRANSYT) e $pcu - s/s$ (descentralizada e fixa).	82
6.2	Proporção de melhoria do algoritmo descentralizado \times TRANSYT.	83
6.3	Proporção de melhoria do algoritmo descentralizado \times tempo fixo.	83
6.4	Atraso (s) - (período de simulação: 20 amostras de tempo) Comparação entre modelos Descentralizado(D) e Global(D).	84
6.5	Proporção de melhoria do modelo global e algoritmo descentralizado.	85

Capítulo 1

Introdução

O tráfego urbano afeta direta ou indiretamente a vida de todas as pessoas que vivem hoje em cidades de grande ou pequeno porte. A circulação de veículos na infraestrutura viária gera conflitos nos deslocamentos, seja por cruzamentos de vias, seja pela travessia de pedestres. Estes fatores podem degradar o desempenho do sistema viário causando atrasos, paradas desnecessárias e acidentes. A sociedade perde com esta situação, pois o tempo em que o usuário fica parado tem sua capacidade produtiva e horas de lazer desperdiçadas, além do *stress* sofrido pelos motoristas, excesso de consumo de combustível e emissão de poluentes.

Os usuários das vias e os administradores que as gerenciam têm a expectativa de um bom funcionamento da rede viária. Este bom funcionamento pode ser avaliado pelo fluxo de veículos contínuo ou com o menor número possível de paradas e atrasos. Os atrasos e paradas ocorrem em sua maioria nos cruzamentos de vias. São também nos cruzamentos que ocorrem em grande parte os acidentes. Assim, surge a necessidade de controlar o fluxo de veículos nestes cruzamentos. Este controle tem como objetivo a solução dos conflitos garantindo segurança aos pedestres e motoristas. O controle pode ocorrer através de regras de prioridade ou por semáforos. As regras de prioridade são baseadas em uma convenção preestabelecida que atribui prioridade de passagem a um determinado conjunto de movimentos do conflito. Os cruzamentos são gerenciados por regras de circulação, sem sinalização específica, previstas no Código de Trânsito Brasileiro [3], sinalizados por placas de vias preferenciais ou parada obrigatória. São funcionais para vias com baixo fluxo veicular onde apresentam melhor desempenho do que o controle semafórico. Para vias de maior fluxo o controle semafórico é o mais utilizado.

O controle de tráfego em interseções realizado por semáforos é hoje a tecnologia mais aplicada em redes urbanas. Torna-se, assim, fundamental garantir um bom ajuste dos semáforos para que a operação viária seja a mais eficiente possível, ao mesmo tempo em que se garante segurança para os usuários.

Uma das estratégias de controle utilizada para ajuste dos tempos semafóricos é chamada de Tempos Fixos. Esta estratégia distribui os tempos semafóricos de acordo com um plano de tempos fixos pré-determinado. Estes planos são elaborados a partir de contagens veiculares para levantamento es-

tatístico das condições de fluxo. Estas contagens permitem a elaboração de planos específicos para diferentes horários do dia e dias da semana [55]. Outro conjunto de estratégias, as de controle atuado, alteram os tempos semaforicos em resposta à demanda verificada através de detectores veiculares e ou pedestres.

A categoria de Controle Atuado é a mais flexível e onde se concentram as pesquisas atuais. Este trabalho encontra-se inserido neste grupo. O objetivo é obter um algoritmo de controle ótimo em tempo real que apresente uma política para solucionar problemas de coordenação semaforica entre interseções e descarga de filas.

1.1 Definição do Problema

Os tempos semaforicos têm um grande efeito sobre o desempenho de uma rede de tráfego. O bom desempenho pode ser avaliado pelo critério de atraso veicular e a garantia de segurança aos usuários das vias. Outra forma de analisar o bom desempenho é através da verificação de uma boa coordenação entre as interseções da rede e a descarga de filas. Para obter-se uma boa coordenação entre interseções, tem-se que garantir que o último veículo da fila (formada durante o tempo de vermelho), liberado da via a montante da interseção (durante o tempo de verde) atingirá a linha de parada enquanto a indicação semaforica, na via a jusante, ainda permanecer verde [31]. Este procedimento deve repetir-se entre as interseções a jusante. Através de uma boa coordenação busca-se reduzir o número de paradas e induzir fluxo contínuo na via arterial. Já o aspecto da descarga de filas é garantido quando a duração do tempo de verde é suficiente para, no mínimo, descarregar a fila formada durante o tempo de vermelho da via, sendo este o caso de fluxo não saturado, ou no limite da saturação. Quando o tempo de verde é suficiente para descarregar também o pelotão que chega da via a montante, ocorre o caso de fluxo não saturado e quando o tempo de verde não é suficiente para descarregar a fila formada no tempo de vermelho do mesmo ciclo, tem-se o caso de fluxo sobre-saturado, o que não será tratado neste trabalho.

Neste trabalho busca-se conseguir uma correta descarga das filas (sem filas residuais após o fim do tempo de verde) com uma boa coordenação através do controle de tráfego em tempo real, ou seja, sem a utilização de planos pré-definidos baseados em estatísticas de contagens de tráfego.

1.2 Objetivo da Pesquisa

Os objetivos deste trabalho podem ser sumarizados como segue:

1. Formular um algoritmo para controle semaforico atuado capaz de induzir coordenação entre interseções e promover fluxo veicular na via arterial;
2. Obter um algoritmo rápido, robusto, no sentido de adequar-se a diferentes situações de demanda veicular e de baixo custo.

1.3 Contribuição da Pesquisa

Para estudar os problemas definidos anteriormente é proposto neste trabalho um modelo para redes de tráfego veicular, obtido a partir do acoplamento de interseções, que permite a análise dos efeitos de diferentes políticas de controle sobre a coordenação de interseções e descarga ótima de veículos. A partir deste modelo foi desenvolvido um algoritmo de controle ótimo em tempo real, com uma arquitetura descentralizada, baseado em simulação. O modelo de tráfego é, inicialmente uma simplificação do modelo de tráfego usado como base, o modelo do PROLYN [21]. As simplificações existem em função das especificidades das redes de tráfego brasileiras. Foram realizadas simplificações também com o objetivo de tornar o modelo e consequentemente o algoritmo mais simples para implementação buscando barateamento de um produto para aplicação nacional.

É apresentada também a formalização matemática do modelo de tráfego em Programação Linear Inteira Mista que permite a utilização de algoritmos e softwares numéricos tais como *Xpress-MP*[33] e *ILOG Cplex* que podem encontrar soluções globais ótimas para o problema de controle semafórico e proporcionar um padrão para comparação de resultados de algoritmos descentralizados.

1.4 Organização do Documento

São apresentados no Capítulo 2 alguns conceitos básicos necessários para compreensão do problema estudado e a solução proposta. A seguir, o Capítulo 3 traz uma revisão dos aspectos que caracterizam e diferenciam as estratégias de controle de tráfego conhecidas e apresenta os modelos de tráfego e as estratégias de controle existentes na literatura. O modelo de tráfego e o algoritmo de controle são apresentados e discutidos no Capítulo 4. No Capítulo 5 é apresentada a proposta de formulação matemática do problema de tráfego para otimização. No Capítulo 6 são apresentados os resultados de simulações para uma rede de 6 interseções sob diferentes condições de fluxo. As conclusões e perspectivas futuras são discutidas no Capítulo 7.

Capítulo 2

Conceitos Básicos de Tráfego Veicular Urbano

2.1 Introdução

Este capítulo apresenta alguns conceitos básicos de tráfego veicular urbano necessários para a compreensão do trabalho. Apresenta também algumas ferramentas e suas classificações para modelar, controlar e otimizar o uso dos sistemas viários.

2.2 Definições Preliminares

O fluxo de tráfego é um processo estocástico que resulta das interações entre veículos, pedestres, condições geométricas de uma via e do controle de tráfego. Nas seções seguintes serão apresentados alguns conceitos necessários para compreensão do comportamento do fluxo de tráfego, suas características e formas de controle. Alguns destes conceitos podem ser encontrados em [1], [9] e [53].

2.2.1 Redes Viárias

As redes viárias são compostas pelos seguintes elementos:

1. **faixa de trânsito** - é o espaço determinado para o fluxo de veículos em um sentido único de fluxo [39];
2. **pista** - é um conjunto de faixas de trânsito;
3. **via** - é um conjunto de pistas que pode permitir sentido duplo de fluxo. Os conceitos de faixa, pista e via são ilustrados na Figura 2.1, para o caso de vias de pista dupla;

4. **interseção** - é o local onde duas ou mais vias se cruzam, criando um conflito entre os sentidos de circulação de veículos;
5. **semáforo** - é um dispositivo de controle de tráfego que alterna o direito de passagem de veículos e pedestres em interseções mediante a utilização de indicadores luminosos. Este sistema de controle organiza de forma cíclica e sequencial a passagem de veículos e pedestres em uma interseção. Existem dois tipos de semáforos, os veiculares e os de pedestres. Eles diferem nas indicações luminosas e, portanto, nas mensagens que transmitem.

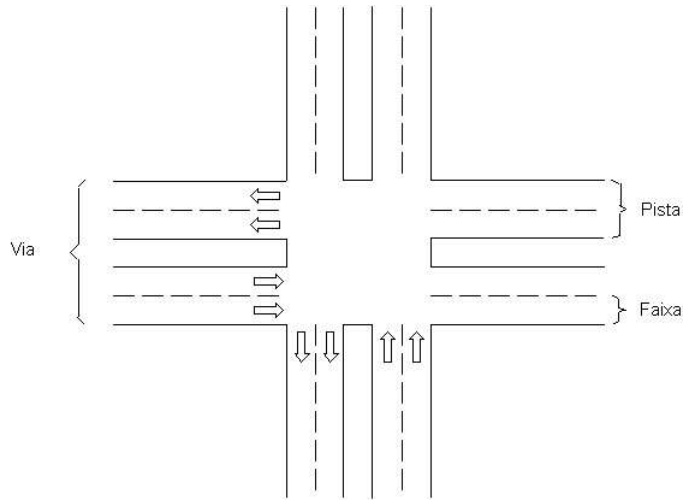


Figura 2.1: Interseção isolada.

2.2.2 Fluxos Veiculares

O fluxo de veículos nas redes viárias pode ser caracterizado pelos conceitos definidos a seguir, que podem ser encontrados em [32], [53]:

1. **velocidade (u)** - expressa a razão de movimento e é usualmente definida como a distância percorrida por unidade de tempo;
2. **velocidade média no tempo (VMt)** - é determinada pela média aritmética das velocidades individuais dos veículos, medidas em um ponto (ou seção) da via (d) durante um determinado intervalo de tempo (t_i);

$$VMt = \frac{1}{n} \sum_{i=1}^n V_i = \frac{\sum_{i=1}^n \frac{d}{t_i}}{n}$$

onde n é o número de veículos e V_i é a velocidade instantânea do i -ésimo veículo.

3. **velocidade média no espaço (VMe)** - é a média das velocidades dos veículos que ocupam um determinado trecho da via durante um intervalo de tempo definido.

$$VMe = \frac{d}{\sum_{i=1}^n \frac{t_i}{n}} = \frac{nd}{\sum_{i=1}^n t_i} = \frac{n}{\sum_{i=1}^n \left(\frac{t_i}{d}\right)}$$

$$VMe = \frac{n}{\sum_{i=1}^n \frac{1}{V_i}}$$

onde VMe é a medida harmônica das velocidades individuais dos veículos.

4. **relação de tráfego contínuo** -

$$q = kVMe$$

onde k é a densidade.

5. **velocidade de cruzeiro** - é considerada como a velocidade que os veículos atingem quando percorrem uma determinada distância sem que ocorram interrupções. Esta velocidade depende das condições geométricas das vias e das condições de tráfego.

6. **fluxo (q)** - é o volume de tráfego expresso em veículos por hora. (Volumes observados por períodos de tempo inferiores a uma hora são geralmente expressos por taxas de fluxo horário equivalente).

$$q = \frac{N}{T}$$

onde N é o número de veículos e T é a unidade de tempo;

7. **volume** - é o número de veículos que passa por uma dada seção de um via em um intervalo de tempo determinado;

8. **densidade ou concentração** - é definido como o número de veículos que ocupam uma unidade de comprimento de uma via em um determinado instante de tempo. Quando a velocidade média (u) e a razão de fluxo (q) são conhecidas a densidade é usualmente calculada como:

$$k = \frac{q}{u}$$

9. **p.c.u.** - é uma unidade que dividida pelo tempo (hora) representa a unidade do fluxo de veículos. A unidade *p.c.u.* é muito utilizada em tráfego e significa *passenger-car units*. Um *p.c.u.* é equivalente a um veículo leve de passeio;

10. **pelotão** - um grupo de veículos que atravessa uma via, sem que ocorra dispersão;

11. **fila** - é um grupo de veículos estacionário na linha de parada de uma via;

12. **headway** - é o tempo que separa a passagem de pontos correspondentes, geralmente o pára-choque dianteiro, de dois veículos consecutivos por uma referência;

13. **headway de saturação** - é o valor de *headway* que corresponde ao fluxo de saturação;

14. **fluxo de saturação (s)** - é a máxima taxa segundo a qual os veículos podem passa por uma dada aproximação se o sinal verde estivesse disponível durante uma hora completa;

15. **taxa de chegada** - é a taxa segundo a qual os veículos chegam em uma determinada faixa de uma pista. Esta taxa pode ser considerada conhecida e constante, ou ainda, medida através de detectores.

16. **taxa de descarga** - é a máxima taxa possível em que uma fila formada em uma faixa da pista é descarregada; esta taxa é igual ao fluxo de saturação enquanto existir fila a ser descarregada;
17. **grau de saturação** - pode ser definido como a relação entre o número médio de veículos que chegam ao cruzamento durante o ciclo através de uma faixa e o número máximo de veículos que podem ser atendidos pelo cruzamento através desta faixa durante um ciclo, sendo dado por:

$$x = \frac{q}{s} * \frac{C}{g}$$

onde: q é o fluxo, s é o fluxo de saturação, C é o comprimento de um ciclo e g é o tempo de verde efetivo.

Portanto, se o grau de saturação for maior que um, significa que chegam mais veículos do que podem ser atendidos naquela faixa. Se esta situação durar por muito tempo, as filas crescem e diz-se que o sistema está saturado;

18. **capacidade da interseção** - é o fluxo total máximo de veículos que pode passar através da interseção em condições operacionais, ou seja, a capacidade não é uma propriedade da interseção propriamente dita, mas de todos os elementos, envolvendo o controle e as condições de tráfego.

2.2.3 Elementos de Controle de tráfego

Uma rede urbana é composta de vias e interseções. Quando estas interseções são semaforizadas, seus tempos semafóricos são ajustados por algum tipo de controle. O fluxo de veículos que circulam nas vias é afetado pelos ajustes realizados pelo controle. Estes ajustes podem ser realizados através de algumas variáveis de controle que serão definidas a seguir, ou ainda pela minimização direta de medidas de desempenho da rede de tráfego, discutidos na seção 2.8. As principais definições para o controle de tráfego são [1], [8], [49], [53]:

1. **tempo de vermelho (r)** - é o tempo durante o qual a luz vermelha do semáforo permanece acesa indicando que os veículos não têm permissão para cruzar a interseção;
2. **tempo de verde (k)** - é o tempo durante o qual a luz verde do semáforo permanece acesa indicando permissão à passagem de veículos;
3. **tempo de amarelo (a)** - é o tempo durante o qual a luz amarela do semáforo permanece acesa indicando que os veículos devem parar antes de cruzar a interseção. Caso não seja possível parar, sem risco para a segurança do tráfego, devem continuar em frente e cruzar a interseção;
4. **tempo perdido (l)** - é a quantidade de tempo dentro de um ciclo que é destinada ao movimentos de uma fase, mas que não é efetivamente aproveitada, devido à aceleração para entrar em movimento quando inicia o tempo de verde e devido à diminuição da taxa de descarga no fim do tempo de amarelo;

5. **tempo de verde efetivo (g)** - é o tempo de verde somado ao tempo de amarelo e subtraído do tempo perdido na fase considerada:

$$g = k + a - l \quad (2.1)$$

6. **ciclo (C)** - é a repetição de uma série básica de combinações semafóricas para uma interseção. Sua duração é chamada de tempo de ciclo. Um exemplo pode ser verificado na Figura 2.2;
7. **fase** - cada conjunto de movimentos comandados por uma mesma sequência de indicações luminosas nos estágios do ciclo;
8. **estágio** - parte de um ciclo durante o qual um grupo de movimentos têm permissão de passagem. Ainda na Figura 2.2 é mostrado um ciclo dividido em três estágios para uma mesma interseção;
9. **entreverde** é o intervalo de tempo entre estágios sucessivos (no qual ocorre a alteração do conjunto de movimentos autorizados e bloqueados);
10. **split** - é a forma como o ciclo está dividido entre os estágios; mais precisamente, é o conjunto de frações do ciclo atribuídas a cada estágio;
11. **offset ou defasagem** - é o instante do início do estágio 1 da interseção, medido em relação a um relógio de referência comum a todas as interseções de um sistema. O *offset* se aplica na sincronização entre interseções que são operadas de forma coordenada como um sistema.

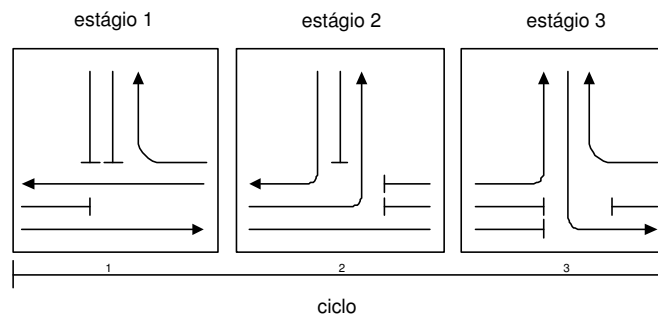


Figura 2.2: Exemplo de um ciclo com sua divisão entre estágios.

2.3 Modelos de Tráfego

Os modelos de tráfego que serão discutidos neste trabalho, em sua maioria, aplicam o conceito de equações de estado para descrever o comportamento do fluxo de tráfego. As equações de estado podem ser descritas por uma estrutura matemática composta de um grupo de n variáveis $x_1(t), x_2(t), \dots, x_i(t), \dots, x_n(t)$ chamadas de *variáveis de estado*, tais que os valores iniciais $x_i(t_0)$ deste grupo e as entradas do sistema $u_j(t), j = 1, \dots, r$ são suficientes para descrever, de forma única, as respostas futuras do sistema para $t \geq t_0$. Existe um grupo mínimo de variáveis de estado para representar o

sistema corretamente. As r entradas, $u_i(t), u_2(t), \dots, u_j(t), \dots, u_r(t)$, são determinísticas, ou seja, elas têm valores específicos para todos os valores no tempo $t \geq t_0$, [17].

O grupo de variáveis de estado $x_i(t)$ representa os elementos ou componentes de um **vetor de estados** n -dimensional $\zeta(t)$:

$$\zeta(t) \equiv \begin{bmatrix} \zeta_1(t) \\ \zeta_2(t) \\ \vdots \\ \zeta_n(t) \end{bmatrix} \equiv \zeta$$

Diversas escolhas para as variáveis de estado podem ser feitas para um sistema de tráfego urbano, sendo algumas delas apresentadas no Capítulo 3.

Quando todas as entradas $u_j(t)$, para um dado sistema, são especificadas para $t \geq t_0$, o vetor de estados resultante determina de forma única o comportamento do sistema para qualquer $t \geq t_0$.

Desta forma, para uma rede de tráfego, definido-se os estados do sistema e a forma como eles evoluem no tempo e, conhecendo-se as entradas do sistema para um período definido, pode-se prever o comportamento futuro da rede de tráfego.

2.4 Tipos de Controle Semafórico

As estratégias de controle que podem ser utilizadas para o ajuste dos tempos semafóricos são classificadas em duas categorias: estratégias de Tempo Fixo e de Controle Atuado [53], [32]. Suas características e sub-divisões serão discutidas a seguir.

2.4.1 Tempo Fixo

As estratégias de controle que trabalham com planos de tempo fixo são baseados em dados históricos da média da demanda de fluxo para um período de tempo em que este plano seja operacional. A elaboração destes planos é realizada através de contagens de veículos. Diferentes planos podem ser gerados para diferentes períodos do dia ou dias da semana para refletir o comportamento estatístico do fluxo de veículos. No entanto, os planos fixos não se adaptam às alterações de demanda que ocorrem em tempo real no dia-a-dia do tráfego. Isto significa que a qualidade do controle de tempo fixo decresce à medida que a variância aumenta, pois quanto maior a variância, maior a dispersão dos dados de demanda histórica e a demanda verificada. Esta dispersão poderia ser corrigida pela atualização dos planos de tempos fixos, mas esta prática envolve custos elevados, deixando de ser realizada por períodos extensos, às vezes anos, implicando na deterioração da otimização dos tempos semafóricos.

Esta obsolescência dos planos semafóricos e a incapacidade de reação a fenômenos de curta duração tem motivado a implantação do controle atuado.

2.4.2 Controle Atuado

As estratégias de controle atuado utilizam-se de técnicas de detecção para monitorar o comportamento do tráfego. Baseados nas informações coletadas são realizados os procedimentos de otimização e implementação de variáveis de controle para o ajuste dos tempos semafóricos.

A categoria de controle atuado pode ser ainda subdividida em:

1. **Seleção dinâmica de planos** - é uma melhoria nas estratégias de tempo fixo em direção às metodologias de controle atuado. É utilizada a monitoração do comportamento de tráfego para definir qual melhor plano de tempos fixos a ser aplicado para a demanda detectada. Este plano é selecionado a partir de uma biblioteca de planos pré-calculados;
2. **Gap crítico** - é uma categoria de controle que permite a mudança de indicação semafórica para verde, em uma via, quando ocorre um espaçamento mínimo entre veículos sucessivos;
3. **Adaptativo** - é a categoria mais avançada de sistemas atuados pelo tráfego, tendo por característica o cálculo *on-line* dos tempos semafóricos, sempre com base em uma estimativa do estado das vias (filas e fluxos) realizada a partir da medição dos fluxos correntes. Eles têm a vantagem de permitir o tratamento de fenômenos dinâmicos de curta duração, evitando que efeitos destes permaneçam por muito tempo nas vias. São em sua maioria descentralizados, determinando seus tempos semafóricos localmente. Os algoritmos propostos neste trabalho situam-se nesta categoria.

2.5 Otimização

A Otimização é a área da Matemática Aplicada que tem como objetivo o cálculo de valores ótimos para variáveis de decisão obedecendo a algum critério de desempenho, ao mesmo tempo que satisfaz restrições de um modelo matemático. A solução de um problema de otimização começa pela transformação do problema em um modelo e, posteriormente a implementação de um algoritmo capaz de encontrar uma solução adequada para o modelo.

Um modelo é uma representação simplificada da realidade que preserva, em determinadas situações e enfoques, uma equivalência adequada. A modelagem de um problema não é uma tarefa trivial, dependendo de fatores subjetivos como intuição, experiência, criatividade e poder de síntese. A formulação de um modelo em linguagem matemática consiste em traduzir o modelo para uma linguagem formal, compreendendo variáveis, equações, desigualdades e fórmulas. Os processos de formulação e validação são iterativos, como mostra a Figura 2.3, pois envolvem múltiplas etapas de tentativa e erro, e interativos à medida que se faz necessária a intervenção contínua do modelador no processo de refinamento do modelo.

A linguagem utilizada para expressar os problemas de maneira declarativa é conhecida como Programação Matemática [60]. Os elementos de um modelo em Programação Matemática são:

1. **Variáveis de decisão** - parâmetros cujos valores definem uma solução para o problema, por exemplo, quantidades produzidas, recursos utilizados ou tempo de verde;
2. **Função objetivo** - uma função das variáveis de decisão, que deve ser maximizada ou minimizada, por exemplo, minimizar custos, reduzir número de homens/hora e maximizar lucros;
3. **Restrições** - um conjunto de funções que define o espaço factível de soluções, por exemplo, limites para recursos, restrições operacionais de um processo de produção, limitações físicas e técnicas.

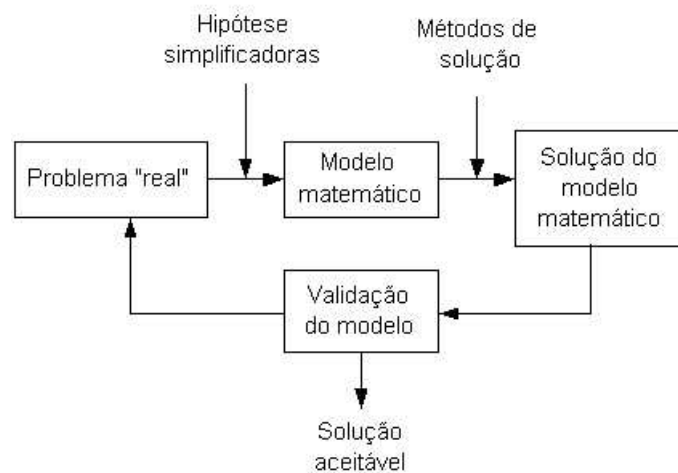


Figura 2.3: Processo de formulação e validação de um modelo matemático.

Um problema geral de otimização pode ser escrito em Programação Matemática como:

$$\text{Minimize} \quad f(x) \quad (2.2)$$

$$\text{Sujeito a :} \quad g(x) \geq 0 \quad (2.3)$$

$$h(x) = 0 \quad (2.4)$$

$$x \in R^n$$

onde $f : R^n \rightarrow R$ é a função objetivo, $g : R^n \rightarrow R^p$ e $h : R^n \rightarrow R^q$ são restrições que limitam o espaço de soluções factíveis, e x é o vetor das variáveis de decisão. Duas exceções a esta formulação geral são problemas sem função objetivo (quando deseja-se apenas encontrar um conjunto de decisões que sejam viáveis) e problemas com múltiplos objetivos.

Dependendo da natureza da função objetivo, das restrições e das variáveis, os problemas de otimização são classificados em subdomínios. A função objetivo para o problema de controle de tráfego urbano, suas restrições, variáveis e o subdomínio de otimização utilizado serão apresentados no Capítulo 5.

2.6 Coordenação entre interseções

Quando interseções semaforizadas são localizadas nas proximidades de outras, o controle de cada interseção isoladamente interfere no comportamento de chegada de veículos nas outras interseções. Estas chegadas deixam de ser randômicas, ocorrendo na forma de pelotões nas interseções a jusante. Os pelotões são criados pela descarga da fila de veículos formada durante a indicação semafórica vermelha. Os pelotões tendem a sofrer dispersão, a qual é função do espaço entre as interseções semaforizadas, do comportamento do motorista e das condições de tráfego ao longo da via. Quando a distância entre interseções não é grande suficiente para que ocorra dispersão, considera-se que o pelotão gerado na interseção a montante chega inalterado na interseção a jusante. Este comportamento motiva a tentativa de coordenar os semáforos de forma a privilegiar a passagem destes pelotões.

O estado ideal de coordenação entre interseções ocorre quando é garantido que o primeiro veículo liberado durante o tempo de verde, atinge o final da fila na via a jusante quando o último veículo desta estiver partindo da linha de retenção. Idealmente, este procedimento deve repetir-se entre as interseções adjacentes. Este fenômeno é ilustrado na Figura 2.4.

A coordenação semafórica, conhecida por *onda verde*, é uma coordenação que ocorre entre os tempos semafóricos com o objetivo de privilegiar a passagem de um dado pelotão. A coordenação que ocorre entre os semáforos não implica, necessariamente, na coordenação dos fluxos de veículos, ou seja, pode ocorrer uma dispersão do pelotão entre as interseções e, conseqüentemente, os veículos não chegarem a tempo de aproveitar o tempo de verde para cruzar a próxima interseção.

Nos algoritmos de Tempo Fixo e nos de Controle Atuado em que a variável *offset* é uma variável de controle, a coordenação semafórica pode ser obtida pelo ajuste desta variável entre interseções vizinhas. Nos algoritmos de Controle Atuado onde o *offset* não aparece explicitamente, e portanto, não é controlado, a coordenação é obtida de forma implícita, ou seja, quando ocorre a troca de informações sobre as chegadas de veículos entre interseções vizinhas durante o processo de otimização. Nos algoritmos de Controle Atuado, a coordenação também pode ser obtida de forma explícita, quando são criados níveis hierárquicos para o controle da coordenação. Um nível superior, que detém informações em nível de rede, define critérios para coordenação global.

Algumas estratégias de controle implementam estas metodologias para obtenção da coordenação. Estas aplicações serão discutidas no Capítulo 3.

2.7 Formação e Descarga de Filas

A evolução determinística das filas pode ser descrita por uma equação de estados onde em um determinado ponto de referência da via, a fila, em um determinado intervalo de tempo, é igual ao número de veículos em fila que existia no tempo anterior, mais o número de veículos que chegam, menos o número de veículos que partem do ponto de referência:

$$x(t) = x(t - 1) + a(t, t - 1) - d(t, t - 1),$$

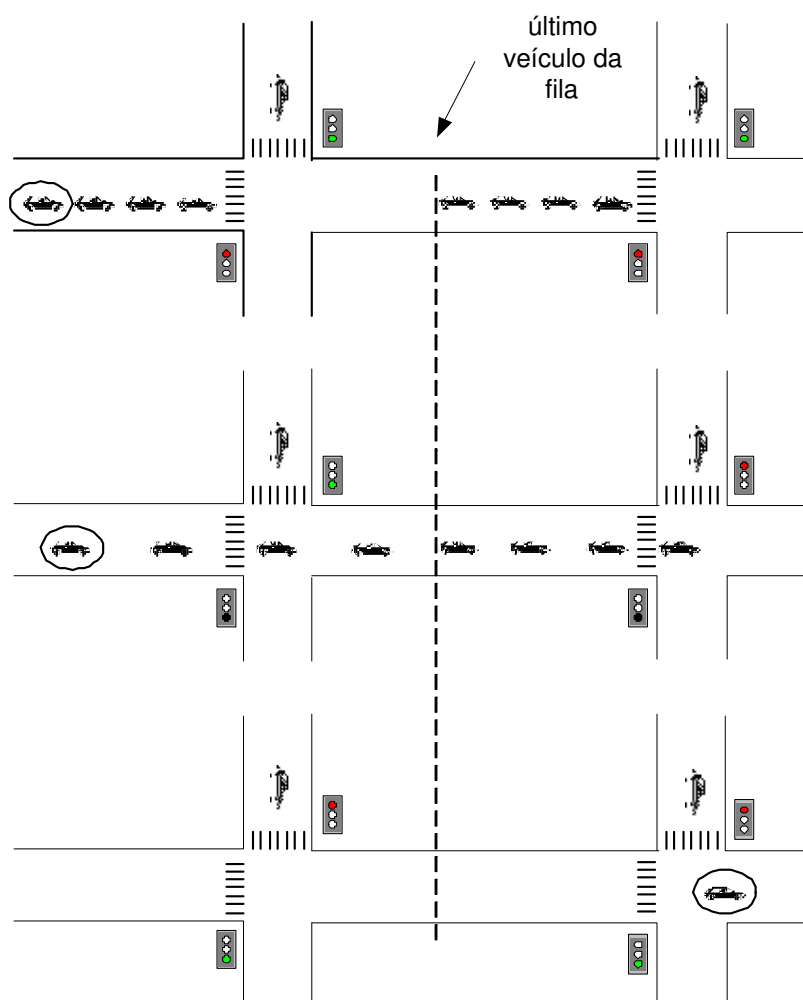


Figura 2.4: Coordenação entre interseções de uma via.

onde $x(t)$ é a quantidade de veículos em fila no tempo t , $a(t, t - 1)$ é o número dos veículos que chegam no intervalo de tempo $(t, t - 1)$ e $d(t, t - 1)$ é o número dos veículos que partem no intervalo de tempo $(t, t - 1)$.

A descarga de filas é realizada a uma taxa máxima de descarga igual ao fluxo de saturação da via. A descarga de fila Q_s pode ocorrer em duas situações [32]: (i) quando a duração do tempo de verde é suficiente para descarregar a fila formada durante o tempo de vermelho da via, este é o caso saturado, o fluxo veicular está no limite da capacidade da interseção, ou seja, o Q_s ocorre no final do tempo de verde; (ii) quando Q_s ocorre antes do término do tempo de verde, o sistema é chamado não-saturado, sobrando tempo para descarregar o pelotão que chega da via a montante. As duas situações descritas, com uma razão de chegadas constante, podem ser avaliadas na Figura 2.5. Uma terceira situação é quando não ocorre o Q_s , ou seja, a fila não é descarregada no tempo de verde, fica uma fila residual para o próximo tempo de vermelho. Este sistema é chamado de sobre-saturado e não será tratado neste trabalho.

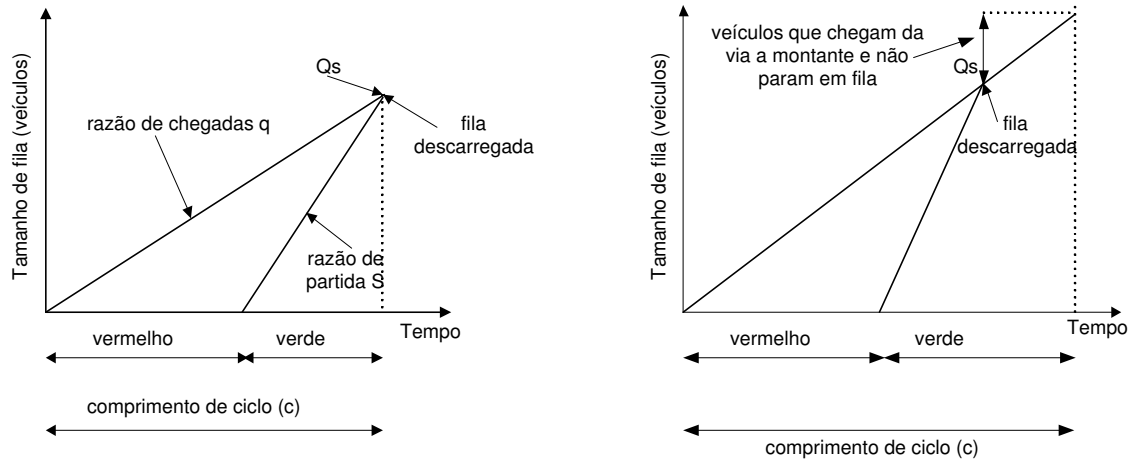


Figura 2.5: Formação e Descarga de Filas.

O fenômeno de formação e descarga de filas pode ser modelado de diferentes formas. Os modelos mais utilizados são o modelo de fila vertical e o modelo de fila horizontal que serão discutidos nas subseções a seguir.

2.7.1 Modelo de Fila Vertical

Vários modelos de fila incorporam duas simplificações críticas para avaliar o atraso dos veículos em interseções semaforizadas. A primeira estabelece a hipótese que os veículos podem desacelerar e acelerar instantaneamente. A Figura 2.6 ilustra o efeito desta suposição nas trajetórias e velocidades dos veículos. Quando os diagramas são analisados, pode ser observado que esta suposição implica na conversão dos atrasos de aceleração e desaceleração em atraso de parada. Assim, a diminuição de velocidade que ocorre nos dois casos, é considerada como se os veículos já estivessem parados. A segunda hipótese assume que os veículos param em fila verticalmente, ou seja, que os veículos atravessam todo o comprimento da via até chegarem na linha de parada, onde formam a fila. Estas

duas hipóteses não interferem na avaliação do atraso, como pode ser visto na Figura 2.6: nos trechos afastados da interseção, as linhas sólida e tracejada são coincidentes. Entretanto, não se pode obter diretamente o comprimento da fila, pois, dada a hipótese de fila vertical, os veículos parecem parar mais tarde do que realmente ocorreria na prática.

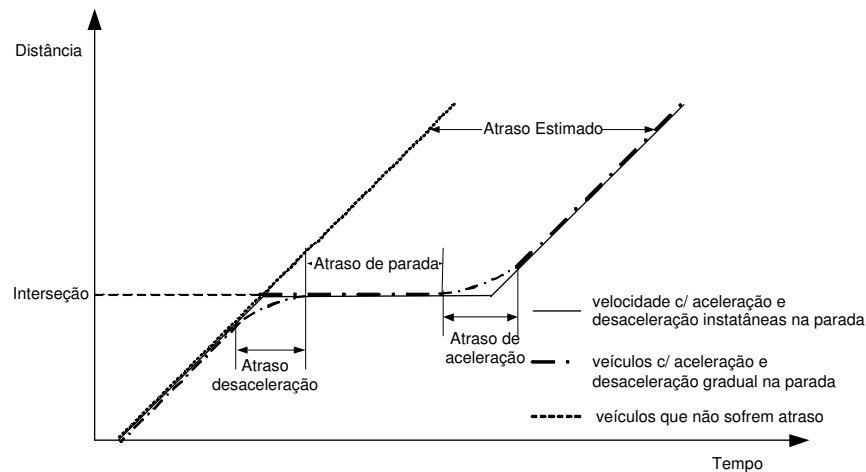


Figura 2.6: Levantamento de trajetória e velocidade com aceleração e desaceleração instantânea

O modelo de fila assume uma razão de chegadas cumulativa e uma razão de partida inicialmente igual a taxa de descarga e, depois de descarregada a fila, igual a razão de chegadas, como pode ser visto na Figura 2.7 (adaptada de [37]). No diagrama, a curva de chegada representa o número de veículos que chegariam na interseção se não sofressem parada no semáforo. A curva de partida, por outro lado, representa o número de veículos que deixam a interseção. Como resultado a distância vertical entre a curva de chegada e a de partida representa o número de veículos que não conseguem atravessar a interseção e portanto, param na linha de parada em fila, enquanto a distância horizontal representa o tempo que um veículo gasta esperando em fila [37].

Ainda na Figura 2.7, três períodos distintos de evolução de tamanho de fila podem ser identificados em um comprimento de ciclo. O primeiro período corresponde a um intervalo durante o qual a curva de partida é horizontal. Este período corresponde a porção de duração do ciclo durante o qual o semáforo tem indicação vermelha e o tráfego não pode atravessar a linha de parada, o que resulta no crescimento do tamanho da fila. O segundo período corresponde a primeira porção da fase de verde durante a qual a fila começa a ser descarregada deixando a interseção a razão do fluxo de saturação. No último período, que somente ocorre em condições de operação não saturada, indica que todas as chegadas atravessam a linha de parada sem sofrerem atraso, ou seja, a fila formada durante o intervalo de vermelho anterior foi completamente dissipada.

Na Figura 2.7, o atraso total no ciclo pode ser estimado pelo cálculo da área entre as curvas de chegada e de partida. Para melhor compreender como o atraso é computado através do diagrama de filas a evolução do tamanho da fila em um ciclo é ilustrado na Figura 2.8. Neste diagrama, pode ser observado que o tamanho máximo da fila ocorre imediatamente antes do início do tempo de verde efetivo. É também observado que o tempo requerido para descarregar totalmente a fila é dado como uma função da diferença entre a razão de chegada em que os veículos chegam ao final da fila e a razão

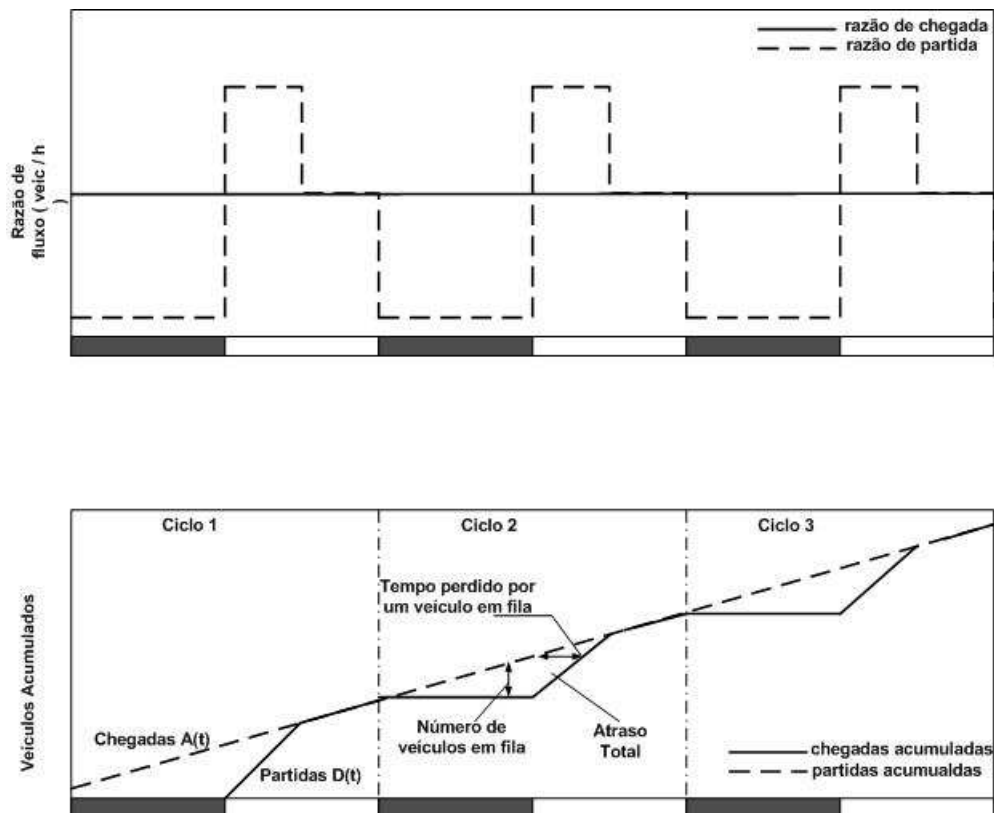


Figura 2.7: Chegadas e Partidas de Veículos Acumuladas para um caso ideal (adaptada de [37]).

de descarga enquanto eles cruzam a linha de parada.

2.7.2 Modelo de Fila Horizontal

Um modelo de fila horizontal, no sentido de considerar a ocupação espacial dos veículos em uma via, foi utilizado e desenvolvido no algoritmo SCOOT [36]. Este modelo será apresentado a seguir.

Como pode ser verificado na Figura 2.9, (adaptada de [56]), um detector é localizado logo após a linha de parada. Desta forma, todos os veículos que chegam na via são detectados. Conhecendo-se o comprimento da via e a velocidade média de percurso dos veículos, torna-se possível determinar o tempo que os veículos vão gastar para chegar na linha de parada da interseção em análise. É necessário definir um tamanho médio para os veículos. Então cria-se duas variáveis: $B(t)$ que define o final da fila, dado em metros e $F(t)$ que define o início da fila, também em metros. O valor inicial para $B(t)$ e $F(t)$ é zero. Quando o semáforo encontrar-se na indicação vermelha, $B(t)$ começa a ter valores crescentes pois os veículos começam a parar um atrás do outro. Depois $F(t)$ começa a crescer (para trás, em distância) quando o semáforo abre e a fila começa a descarregar. Seus valores são crescentes até encontrar o final da fila, e então, $B(t) = F(t)$, quando as duas são levadas a zero. Isto significa que os veículos que continuarem chegando não terão mais que parar em fila.

Desta forma, consegue-se modelar o comportamento de formação de filas bem como a chegada de veículos na fila em diferentes pontos da via.

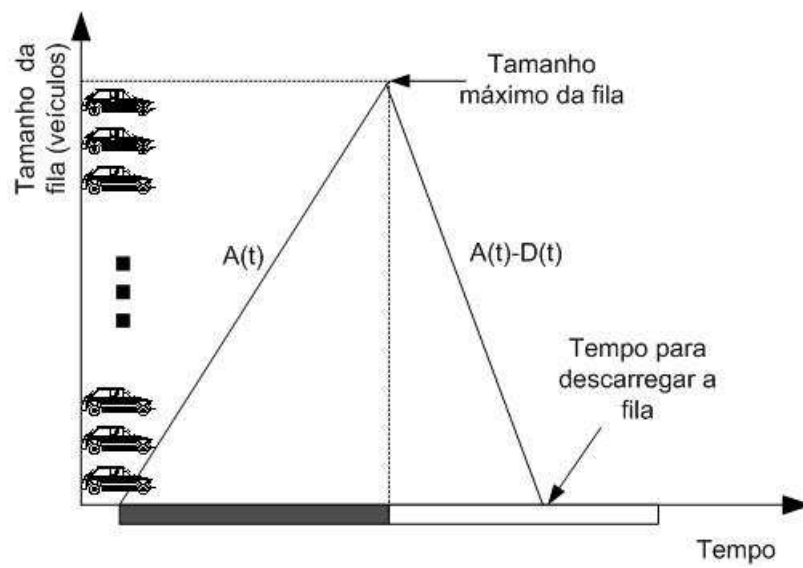


Figura 2.8: Modelo de fila vertical.

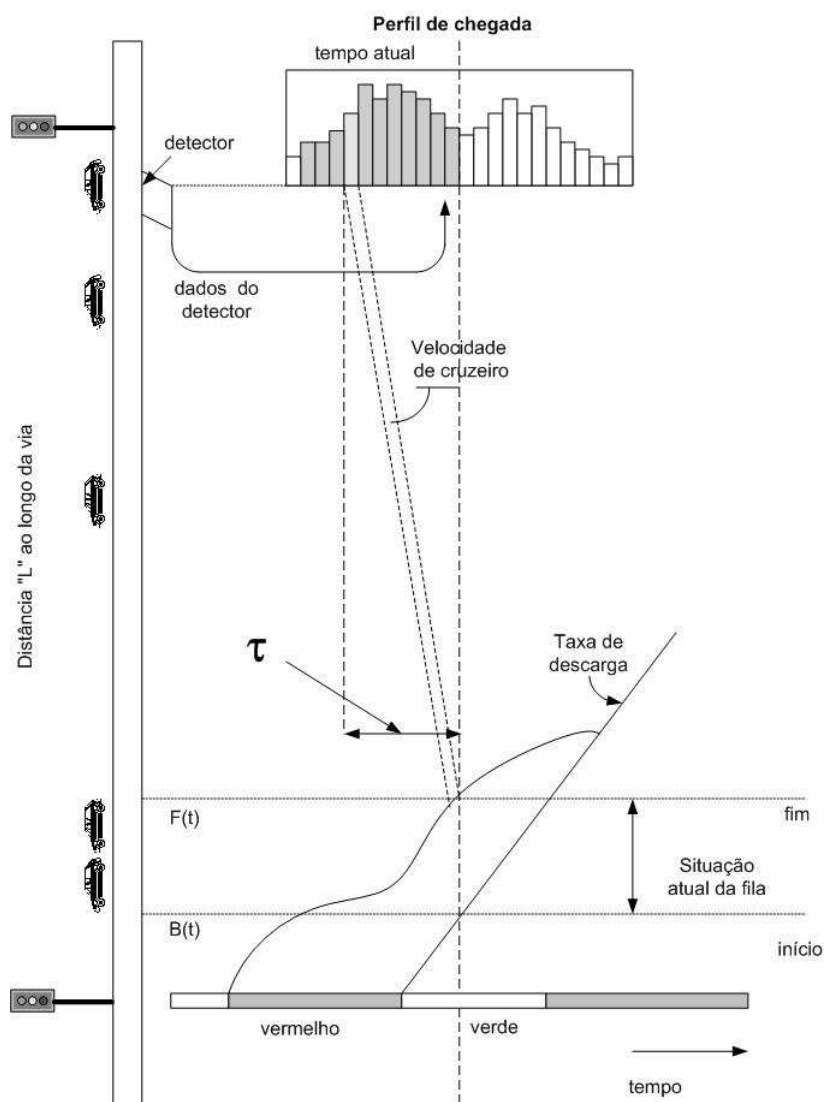


Figura 2.9: Fila horizontal (adaptada de [56]).

2.7.3 Comparação entre os modelos de fila

O modelo de fila vertical é equivalente ao modelo de fila horizontal, para o caso de sistema não saturado. Isto pode ser ilustrado na Figura 2.10. Observa-se, na Figura, que as linhas de A a B são as mesmas fora da região sombreada. Esta equivalência é esperada pois os dois modelos devem representar o evento fila de forma a reproduzir seu comportamento real. A principal diferença entre os dois está no fato do modelo horizontal, ao modelar a ocupação espacial dos veículos, permitir a obtenção com maior precisão e de forma explícita a localização do último veículo em fila, proporcionando melhores condições para modelagem da coordenação dos fluxos e para previsão de bloqueios de vias a montante.

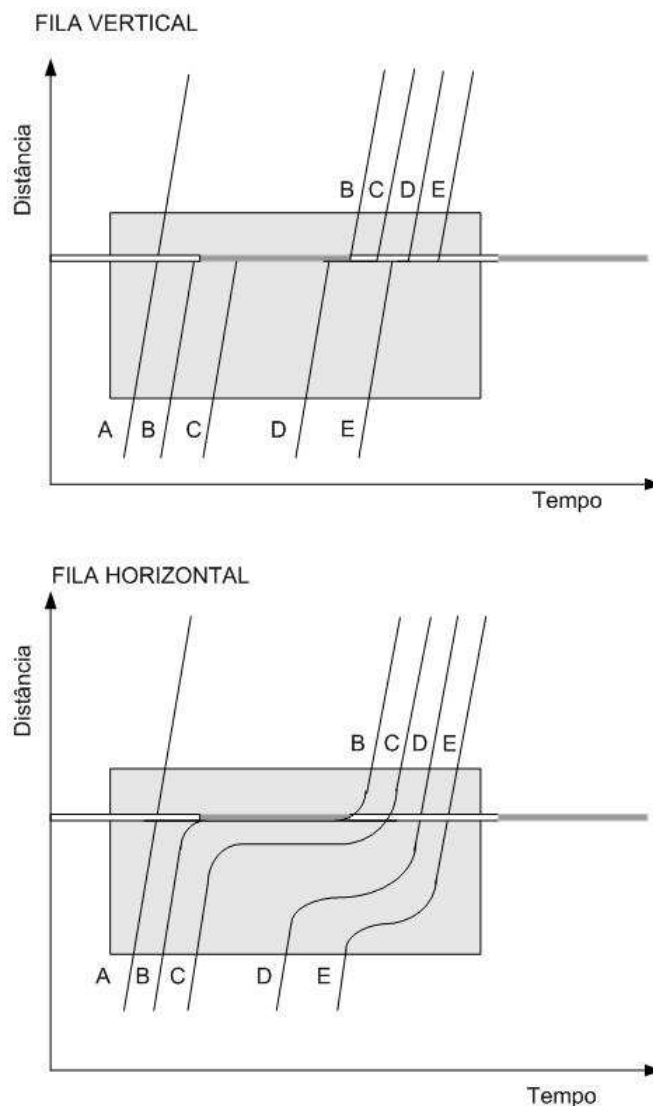


Figura 2.10: Comparação entre modelo de fila horizontal e vertical.

2.8 Critério de Desempenho

O desempenho de um sistema viário pode ser medido através de uma combinação de critérios: atraso acumulado de veículos, número de paradas, consumo de combustível e poluição atmosférica. Os critérios de consumo de combustível e poluição atmosférica são aspectos que surgem como uma consequência dos atrasos e paradas de veículos, pois os picos de consumo e poluição ocorrem justamente nas interseções semaforizadas onde os veículos são obrigados a desacelerar, parar e depois acelerar novamente para entrar em movimento.

2.8.1 Atraso

O critério de atraso veicular acumulado está diretamente relacionado à formação de filas. Ele avalia o desempenho do tráfego através do número de veículos parados em fila. Assim quanto maior a fila, maior o atraso. A formação de filas, como discutido na seção 2.7 ocorre nas interseções em duas situações: (i) durante o vermelho, quando os veículos são obrigados a parar, (ii) quando a razão de chegadas é maior do que a capacidade de descarga da via durante o tempo de verde do semáforo. O último caso ocorre em situações de sobre-saturação [58].

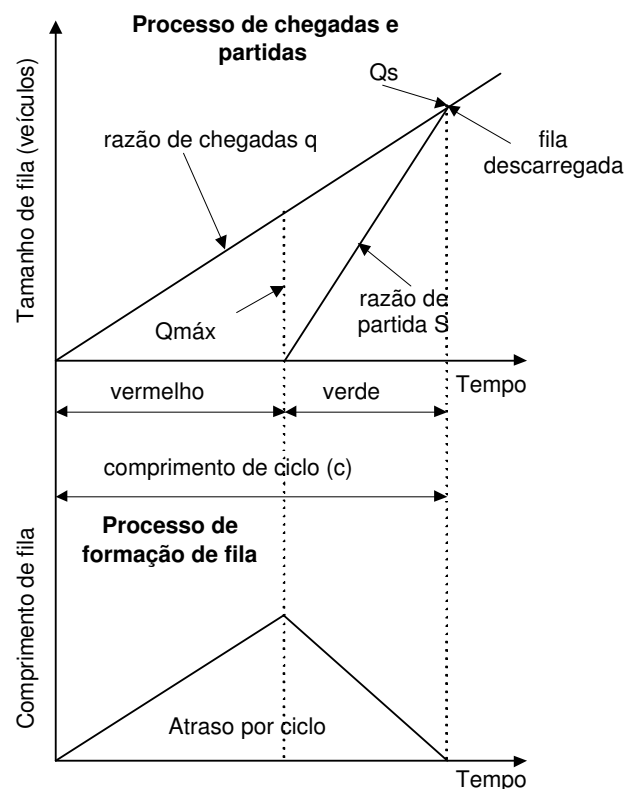


Figura 2.11: Componentes da formação do atraso (adaptado de [58].)

A situação (i) pode ser observada na Figura 2.11, (adaptado de [58]). Considera-se uma razão de chegadas constante e fila inicial igual a zero no início da fase vermelha. Neste caso é considerado um fluxo de chegadas constante e uniforme q . Com a mudança de indicação semaforica para verde,

e tendo-se ainda chegada de veículos, a fila continua crescendo a mesma razão q . A distribuição de partidas uniforme é realizada a uma razão de fluxo de saturação constante S , enquanto existir a fila. Se a fila formada durante o tempo de vermelho for totalmente descarregada antes de terminar o tempo de verde, (ponto Q_s) os veículos que continuarem chegando passam em velocidade de cruzeiro pela via sem precisar parar em fila. A partir de Q_s , a razão de partida dos veículos é igual a razão de fluxo de chegada. Na parte inferior da Figura 2.11 é representada uma composição do carregamento e descarga da fila, onde pode ser verificado o tamanho da fila crescendo e decrescendo. A área sob a curva determina o atraso ocorrido durante o ciclo avaliado.

A situação (ii) pode ser observada na Figura 2.12. A razão de chegadas maior do que a capacidade de descarga da via provoca o acúmulo de veículos em fila. A área sob o atraso uniforme torna-se crescente, caracterizando a sobre-saturação, ou seja, o bloqueio de filas que não podem ser descarregadas em um mesmo ciclo.

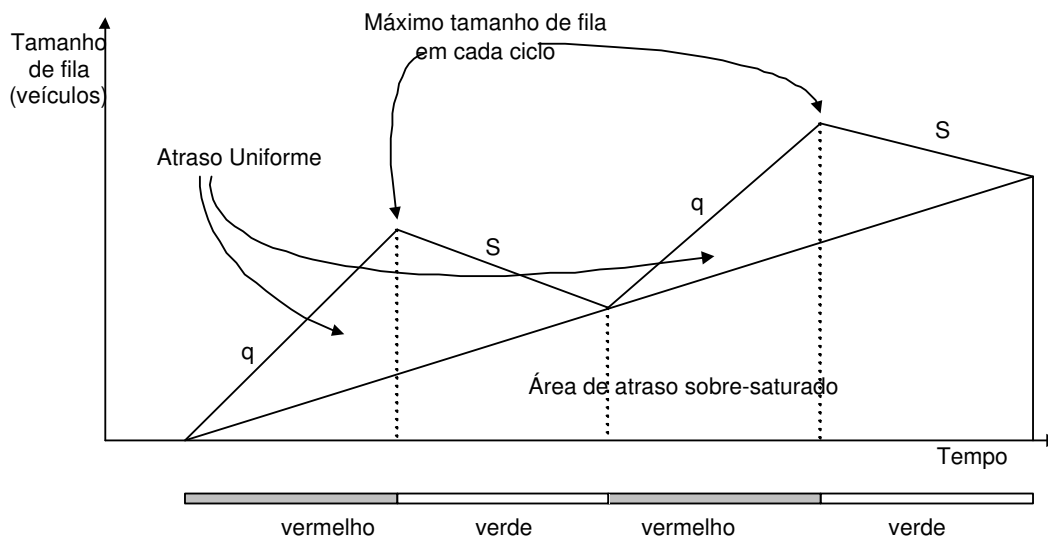


Figura 2.12: Componentes da formação do atraso sobre-saturado.

O processo de cálculo do atraso depende do conhecimento da variável número de veículos em fila. Para quantificá-la é necessário conhecer o fluxo de veículos nas vias. Este fluxo pode ser obtido diretamente das contagens realizadas por sensores localizados nas vias ou manualmente por levantamento de dados históricos. Dependendo do arranjo dos sensores é possível obter o tamanho das filas ou apenas a quantidade de veículos em movimento ou ainda, a sinalização de algum veículo parado sobre um sensor. Além disso, pode-se desejar instalar um número menor de sensores do que o necessário para cobrir todas as faixas de rodagem, por questões de economia. Neste caso, a estimação das filas deverá ser feita com informação parcial dos fluxos veiculares. A previsão das filas, ou seja, sua quantificação, e o atraso médio acumulado no horizonte serão discutidos no Capítulo 4.

Nos algoritmos de tempo fixo e os de controle atuado com definição da variável *offset*, o critério de atraso está relacionado com esta variável, pois um ajuste sincronizado dos semáforos adjacentes pode proporcionar uma diminuição no tamanho da fila em uma via arterial.

2.8.2 Número de Paradas

O critério de desempenho número de paradas é obtido pela avaliação da quantidade de todo tráfego que sofre atraso, ou seja, quando tem-se veículos parados em fila, existe contribuição para o número de paradas. No entanto, diferentes situações onde ocorre uma parada de veículo devem ser consideradas:

- quando a fila está sob indicação semafórica vermelha - todos os veículos parados em fila contribuem para o número de paradas;
- quando os veículos encontram o final da fila e está já recebeu indicação semafórica verde - os veículos que por ventura, pararem ao encontrar o final da fila ainda parada e os veículos que não conseguirem cruzar a interseção dentro da indicação semafórica verde, contribuem para o número de paradas;
- quando a fila existente é descarregada, mas o veículo recebe indicação semafórica vermelha ao chegar próximo a linha de parada neste caso os veículos que não conseguem atravessar a interseção contribuem para o número de paradas.

As situações descrita anteriormente podem ser descritas matematicamente da seguinte forma:

$$n_p^l(t+1) = \begin{cases} a_{l,1}(t) & \text{se } x_l(t) \geq s_l m_l(t) \\ x_l(t) + a_{l,1}(t) - s_l m_l(t) & \text{se } x_l(t) < s_l m_l(t) \text{ e } x_l(t) + a_{l,1}(t) > s_l \\ 0 & \text{se } x_l(t) < s_l m_l(t) \text{ e } x_l(t) + a_{l,1}(t) \leq s_l \end{cases}$$

onde $n_p^l(t)$ é a variável que indica o número de paradas para cada via l ; $a_{l,1}(t)$ representa o número de veículos que chegam na interseção e já percorreram toda a extensão da via, encontrando-se próximo a linha de parada; $x_l(t)$ representa a quantidade de veículos em fila; s_l é o fluxo de saturação e $m_l(t)$ é a variável de controle que indica se o semáforo está com indicação verde ou vermelha.

2.8.3 Comparação entre os critérios atraso e paradas

Será possível verificar no Capítulo 3 que os critérios de atraso e paradas são utilizados ora em conjunto, ora apenas o de atraso. Nesta seção será discutida esta questão.

Dois exemplos para auxiliar na compreensão do peso dos critérios serão apresentados. A situação será exemplificada com o auxílio da Figura 2.13. Tem-se um pelotão de veículos aproximando-se da interseção com seu tempo de aproximação do semáforo de 10s. O tempo de verde mínimo considerado é de 12s. Na via em conflito tem-se apenas um veículo chegando a, por exemplo, 2s do semáforo. V_1 aguarda a abertura do semáforo. Duas opções podem ser consideradas:

- abre-se o semáforo para V_1 . Neste caso, o atraso resultante é de 2s para cada H_i , $i = 1, \dots, 4$, pois estes 2s (tempo de verde mínimo - tempo de percurso) são o tempo de parada devido ao verde mínimo da interseção;

- V_1 aguarda a passagem do pelotão. Neste caso, o atraso é maior do que 10s. Portanto, a decisão (i) é tomada, e o pelotão é forçado a parar. Em suma, o sinal verde para a via vertical minimiza o atraso, porém força a parada de um pelotão na via horizontal.

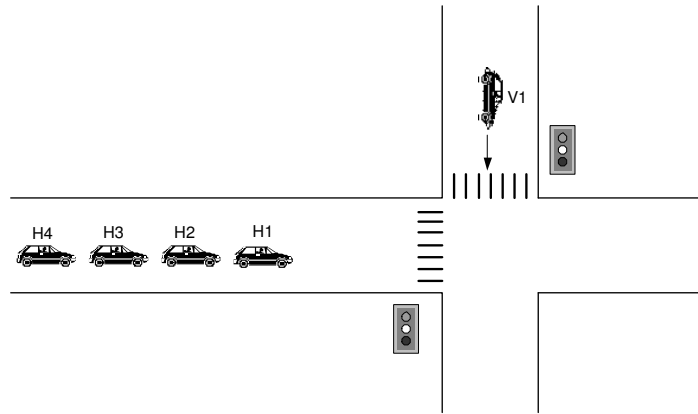


Figura 2.13: Exemplo de aplicação do critério de atraso.

Para o critério de número de paradas aplicado isoladamente, tem-se o seguinte exemplo (Figura 2.14): supõe-se 6 veículos parados na interseção, durante a indicação vermelha. Na via em conflito estão chegando dois veículos, um distante do outro. No entanto, pelo critério de parada a indicação verde permanecerá na via com menor número de veículos, pois estes ainda não pararam. Isto ocorrerá até que o limite máximo de verde seja atingido, independentemente da retenção do pelotão na outra via.

Apesar de, conforme visto na Figura 2.13, o uso do critério de atraso puro poder levar a um número maior de paradas do que se fossem ponderadas estas últimas, a dificuldade em se encontrar um bom valor de ponderação leva, comumente, à adoção do atraso como critério único (um exemplo de ponderação pode ser visto em [13]).

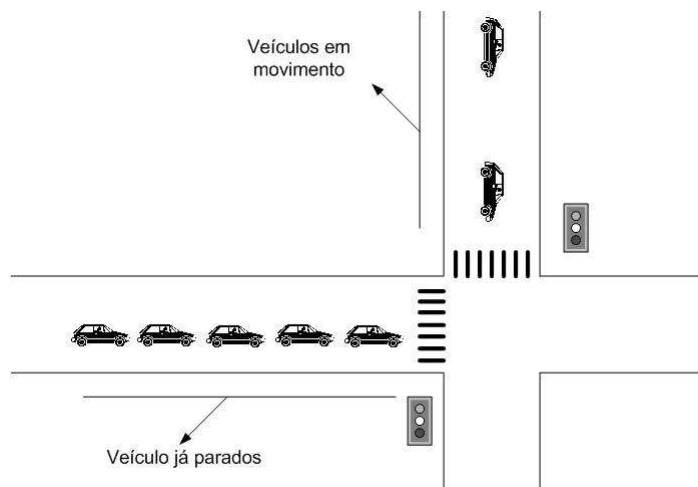


Figura 2.14: Exemplo de aplicação do critério de parada.

2.9 Conclusões

Neste capítulo fez-se uma revisão de alguns conceitos básicos na descrição de um sistema de tráfego urbano. Discutiu-se também algumas ferramentas para modelar, controlar e otimizar as redes de tráfego urbanas. Por fim, as formas de se avaliar o desempenho das redes de tráfego e características como a coordenação e descarga de filas também foram discutidas.

Capítulo 3

Revisão Bibliográfica

3.1 Introdução

No Capítulo 2, foram apresentadas as duas categorias em que pode-se dividir as estratégias de controle de tráfego: Tempos Fixos e Controle Atuado. Neste Capítulo serão apresentados alguns modelos de tráfego existentes na literatura e suas estratégias de controle, em especial o modelo de tráfego do PRODYN [21], que foi utilizado como base para o desenvolvimento do modelo utilizado neste trabalho.

Será apresentada também uma classificação dos modelos de tráfego e suas estratégias de controle através do agrupamento de suas características.

Deseja-se também, guiar o leitor na discussão de como as estratégias apresentadas tratam os problemas de coordenação e descarga de filas, reservando, para o Capítulo 4, a discussão comparativa entre o algoritmo proposto e os existentes.

3.2 Modelos de Tráfego e Estratégias de Controle Associadas

A necessidade de solucionar os conflitos do tráfego urbano despertaram interesse de pesquisa desde meados da década de 50. Inicialmente os trabalhos concentraram-se na determinação do atraso em interseções isoladas, entre eles Webster [59]. O caminho trilhado por Webster e por vários outros pesquisadores foi o de tentar estabelecer uma relação matemática entre o atraso, as variáveis de temporização do controlador, a demanda de tráfego e ainda o fluxo de saturação na região da interseção.

O grau de saturação é uma variável que também depende dos tempos semafóricos, sendo definido como a relação entre o número médio de veículos que chegam ao cruzamento durante o ciclo através de uma faixa e o número máximo de veículos que podem ser atendidos pelo cruzamento através desta faixa durante um ciclo. Portanto, se o grau de saturação for maior que um, significa que chegam mais veículos do que podem ser atendidos naquela faixa. Se esta situação durar muito tempo, as filas crescem e diz-se que o sistema está saturado.

O percentual de verde efetivo (λ) que é atribuído a uma faixa de uma pista definido como:

$$\lambda = \frac{g}{C} \quad (3.1)$$

onde C é o tempo de duração do ciclo e g é o verde efetivo. Da definição de *grau de saturação* e de (3.1), pode-se chegar a:

$$x = \frac{Cq}{gs} = \frac{Cq}{\lambda Cs} = \frac{q}{\lambda s}$$

onde x é o grau de saturação, q é o fluxo observado e s o fluxo de saturação da faixa considerada.

A relação entre a demanda de fluxo, as variáveis de temporização e o grau de saturação estabelecidos por Webster é dada por:

$$d = \frac{C(1-\lambda)^2}{2(1-\lambda x)} + \frac{x^2}{2q(1-x)} - 0.65 \left(\frac{C}{q^2} \right)^{\frac{1}{3}} x^{(2+5\lambda)} \quad (3.2)$$

onde:

- d é o valor atraso médio total por veículo (segundos/veículos);
- λ é a proporção do ciclo efetivamente aproveitada pelos veículos para se movimentarem através da interseção (verde efetivo) (g/C);
- C é a duração do ciclo (segundos);
- q é o fluxo observado (veículos/segundos);
- g é o tempo de verde efetivo (segundos);
- x é o grau de saturação.

Na equação (3.2) o primeiro termo representa a média do atraso dos veículos de chegada uniforme. O segundo termo estima o atraso adicional devido às chegadas randômicas de veículos. Este atraso adicional é atribuído à probabilidade da chegada de veículos causarem uma sobre-saturação temporária. O terceiro termo, finalmente, é um fator de ajuste que foi introduzido ao modelo para corrigir o atraso estimado, tendo sido obtido empiricamente, a partir de estudos de simulação.

A Figura 3.1, apresenta o resultado gráfico da equação (3.2), quando o atraso médio por veículo para uma faixa da pista é plotado contra o fluxo q , para um fluxo de saturação de 1800 v.p.h. e um percentual de verde efetivo de 50%. Observa-se que o atraso tende ao infinito quando o fluxo ultrapassa um certo limite, independentemente do tamanho do ciclo. Isto se dá quando o grau de saturação tende a 1.

No entanto, o resultado mais efetivamente aplicado com este desenvolvimento é o cálculo dos tempos semaforicos para estratégias de tempo fixo, através da minimização do atraso calculado pela fórmula proposta por Webster. Estes cálculos de temporização serão apresentados logo após a apresentação da fórmula do atraso. A temporização de um sistema semaforizado que opera com ciclo fixo

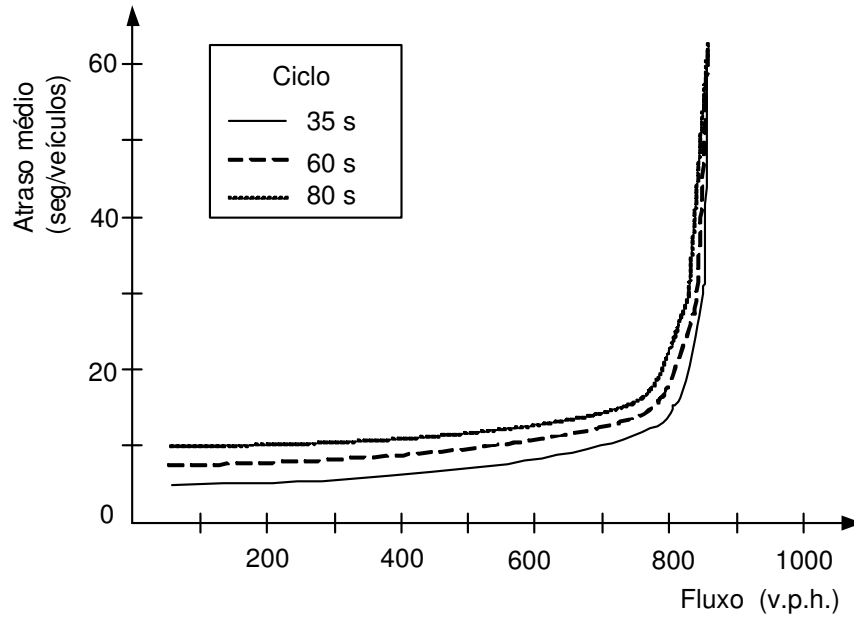


Figura 3.1: Atraso médio veicular de uma faixa

determina basicamente dois fatores: o ciclo que fornece o mínimo atraso e a divisão deste ciclo entre diversas fases (*split*).

O cálculo dá-se inicialmente pela determinação das taxas de ocupação das faixas que chegam em uma interseção. Seleciona-se a máxima taxa de ocupação (Y_i), que é a relação entre o fluxo observado q e o fluxo de saturação s , dentro de cada grupo de movimentos e distribui-se o tempo total de verde efetivo para cada fase na proporção destas máximas relações encontradas:

$$Y_i = \max_j(y_{i,j})$$

onde $y_{i,j}$ é a taxa de ocupação de cada faixa j de uma via i .

A faixa de cada grupo de movimento que possui o maior $y_{i,j}$ é chamada de *faixa crítica* do grupo. No exemplo da Figura 3.2:

Para o movimento N-S as relações q/s são:

$$y_{1,1} = 600/2000 = 0,30$$

$$y_{1,2} = 600/2400 = 0,25$$

$$y_{1,3} = 600/3000 = 0,20$$

$$y_{1,4} = 600/3000 = 0,20$$

$$Y_1 = \max(y_{1,1}, y_{1,2}, y_{1,3}, y_{1,4}) = 0,30$$

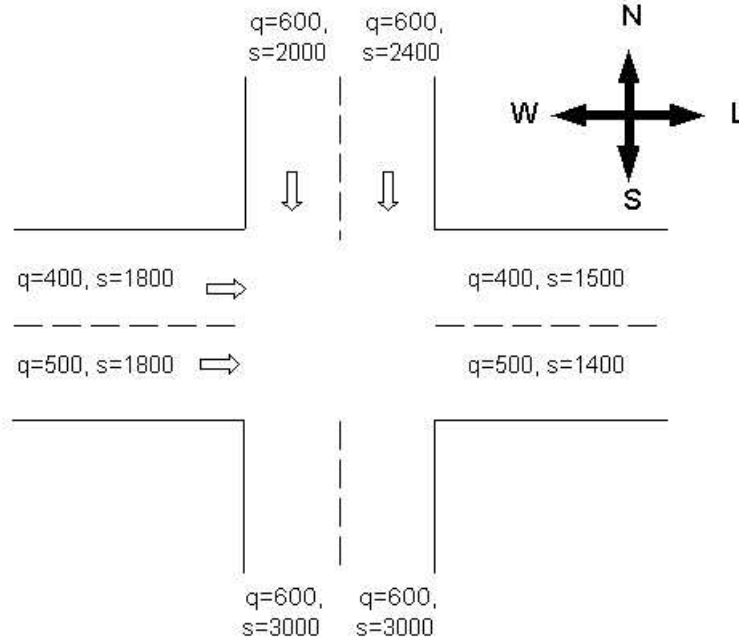


Figura 3.2: Exemplo para o cálculo do *split* [12]

Para o movimento W-L:

$$y_{1,1} = 400/1800 = 0,22$$

$$y_{1,2} = 500/1800 = 0,28$$

$$y_{1,3} = 400/1500 = 0,27$$

$$y_{1,4} = 500/1400 = 0,36$$

$$Y_1 = \max(y_{2,1}, y_{2,2}, y_{2,3}, y_{2,4}) = 0,36$$

Portanto, o tempo de verde efetivo deverá ser dividido na proporção de 0,30/0,36 para os movimentos N-S e W-L.

Dados o número de fases n , o tempo de vermelho de segurança por fase R e o tempo perdido por fase l , o tempo total L perdido durante um ciclo é dado por:

$$L = n(l + R)$$

O verde efetivo total para o ciclo (g) torna-se:

$$g = C - L$$

Para determinar o tempo de verde efetivo para uma fase i (g_i) faz-se:

$$g_i = g \frac{Y_i}{\sum_{j=1}^n Y_j} \quad (3.3)$$

Finalmente, pode-se determinar o tempo de verde associado a cada fase (k_i) através da equação (2.1), reescrita como:

$$k_i = g_i - a_i + 1 \quad (3.4)$$

Para determinação do ciclo ótimo Webster utilizou a equação (3.2) com modificações para chegar a equação:

$$C = \frac{1,5L + 5}{1 - \sum_{i=1}^n Y_i} \quad (3.5)$$

As unidades de C e L são em segundos. Esta expressão fornece a duração de ciclo que leva ao mínimo atraso total por veículo para interseções isoladas semaforizadas e com tempo fixo. Para chegar até ela, a equação (3.2) foi manipulada para obter uma expressão para o atraso total médio de todos os veículos que cruzam a interseção. Em seguida, a equação resultante foi minimizada em relação ao ciclo.

Após o trabalho de Webster, numerosos estudos foram desenvolvidos tendo como objetivo a estimação do atraso em interseções isoladas semaforizadas. Como resultado destes estudos, um número de modelos de atraso baseados na teoria de fila foram propostos para diferentes situações de tráfego. Dentre estes modelos pode-se citar o desenvolvido por Miller [5] e Akcelik [6] e os modelos desenvolvidos para uso no *Highway Capacity Manual* [8].

O passo seguinte na literatura foi o desenvolvimento de modelos que controlassem interseções semaforizadas ao longo de uma via arterial, considerando assim o conjunto de semáforos como um sistema. O conceito de controle de fluxo de tráfego em uma arterial pode ser descrito graficamente como mostrado na Figura 3.3. A Figura introduz a necessidade das seguintes definições:

1. *Banda verde* é o espaço entre um par de linhas paralelas que representam velocidades que delineiam um movimento progressivo no espaço e no tempo;
2. *Velocidade de progresso* é a inclinação da curva banda verde que representa a velocidade dos veículos ao longo da banda verde;
3. *Largura da banda* é a largura de uma banda verde em segundos. Indica o período de tempo disponível para o fluxo de tráfego em uma banda.

O controle de fluxo de tráfego em uma arterial, a partir dos conceitos apresentados, parte do princípio que, se um determinado número de usuários mantiverem seus veículos em uma dada velocidade constante, estes encontrarão uma sequência de tempos de verde que minimizará o atraso e as paradas deste grupo de veículos. Isto porque os três conceitos enumerados anteriormente determinam uma velocidade, um tempo e o espaço de percurso deste veículos para garantir uma banda de passagem com minimização de atrasos.

Os modos de operação definidos, a velocidade de progressão e a largura de banda foram a base para o desenvolvimento de planos de tempos semaforicos fixos, com o objetivo de maximizar a *onda verde* nas arteriais [61]. Está dentro desta categoria o algoritmo MAXBAND [40], onde um arranjo

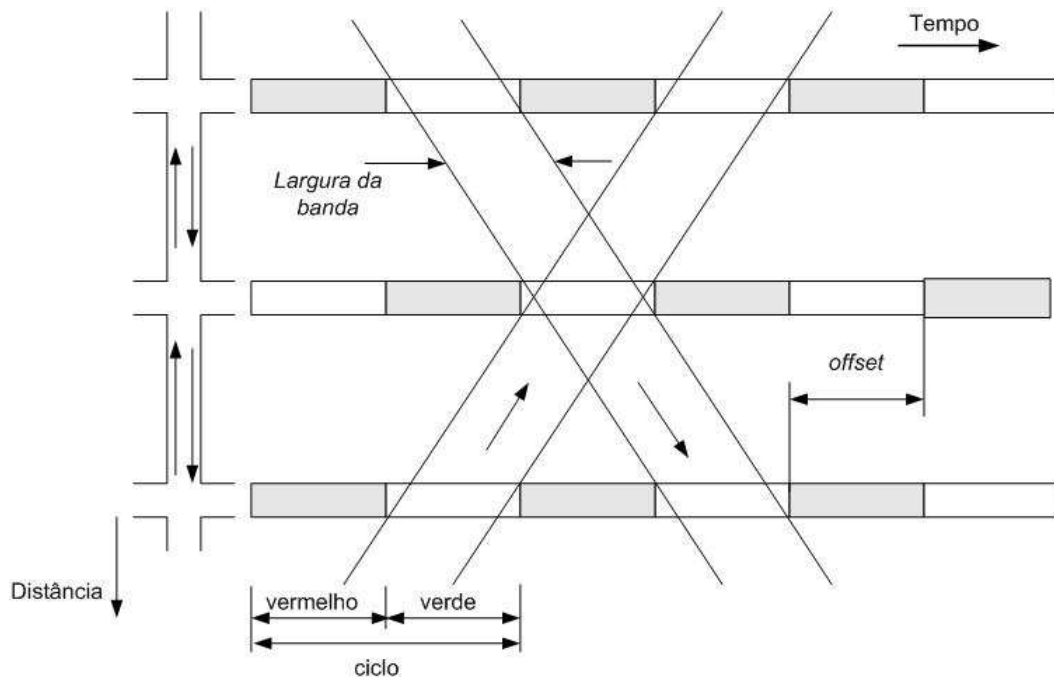


Figura 3.3: Diagrama espaço-tempo

dos tempos de verde é realizado com o objetivo de maximizar as bandas para passagem de pelotões ao longo de uma via arterial. Note que o objetivo é, dados os tempos de verde, obter a máxima largura de banda, independentemente dos fluxos veiculares. Assim, pode ocorrer de um bom cálculo de *onda verde* entre semáforos resultar em progressão ruim devido a filas formadas durante o vermelho nas interseções a jusante. Já programas de coordenação como o TRANSYT calculam o *offset* considerando os fluxos previstos para a malha viária, (isto é, levam em conta as filas formadas no vermelho) de modo a minimizar, principalmente, o atraso médio veicular e o número de paradas.

Na década de 70, surgiram algoritmos de controle atuado, que permitem que os tempos semafóricos respondam as mudanças ocorridas nas condições de tráfego em tempo real. Inicialmente, surgiram algoritmos que utilizam os conceitos de *split*, *offset* e duração de ciclo, calculados ou incrementados a partir de informações recebidas de detectores veiculares, como o SCATS [42], e o SCOOT [36]. Durante as duas décadas seguintes surgiram os algoritmos que buscam a minimização de critérios de desempenho da rede através de algoritmos de otimização. Pode-se citar alguns deles como PROLYN [21], OPAC [27], RHODES [45], ALLONS-D [52] e CRONOS [10].

As estratégias de controle de tráfego citadas serão apresentadas nas seções seguintes.

3.2.1 TRANSYT

TRANSYT, *Traffic Network Study Tool*, será o primeiro modelo apresentado sendo o principal representante da categoria de tempos fixos [55]. O TRANSYT permite a coordenação dos tempos semafóricos entre as interseções através da otimização de *offsets* entre os grupos de semáforos adjacentes, levando em conta os fluxos que ocorrem na malha viária. Os planos de tempos fixos especi-

ficam as variáveis de controle: duração da fase, o tempo de ciclo e os *offsets* para cada interseção na rede. Estes são implementados em cada central localizada nos respectivos semáforos definindo para cada interseção quando deve ocorrer sua mudança de fase. Várias são as versões que existem hoje do TRANSYT, as principais linhas que surgiram foram a versão britânica, TRANSYT/10, versão utilizada para comparação e análise neste trabalho, e a versão americana TRANSYT/7F.

O princípio de funcionamento e a estrutura básica da estratégia de controle TRANSYT será discutida com o auxílio da Figura 3.4.

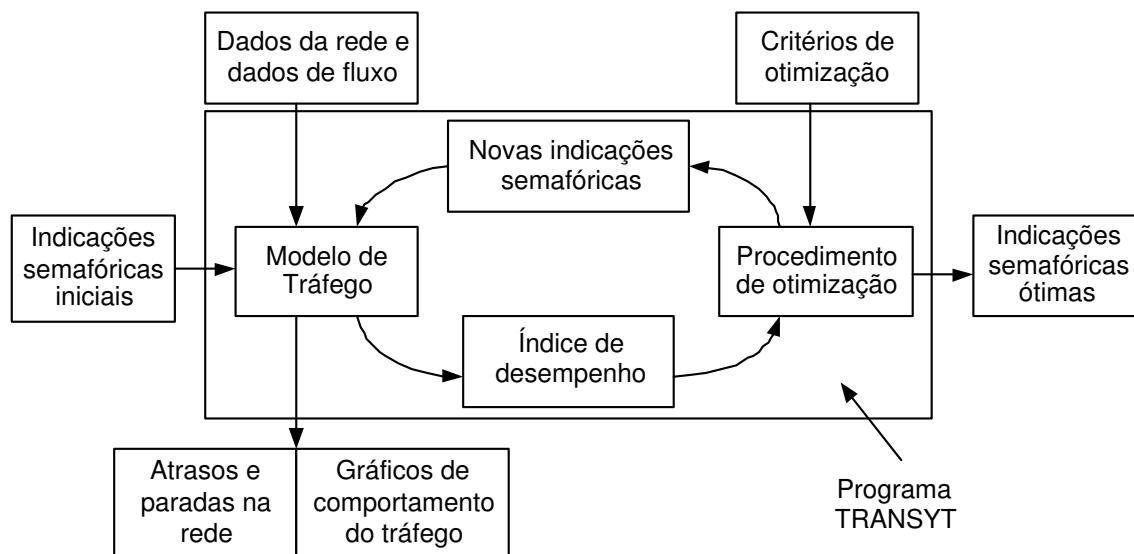


Figura 3.4: Modelo TRANSYT

As variáveis de entrada do TRANSYT são as condições semaforicas iniciais, incluindo o estágio pré-especificado, a duração de verde mínimo para cada estágio e cada interseção, e os valores iniciais de *offset*, *split* e ciclo. Os dados de rede e fluxo compreendem: fluxos de saturação, tempos de viagem das vias, taxas de conversão (constantes e conhecidas) para cada interseção e as demandas também constantes e conhecidas. Condições de sobre-saturação, quando a capacidade da via é ultrapassada, não são descritas na versão utilizada. Em versões mais recentes, é possível definir o número máximo de veículos por *link* que, quando excedido, afeta o índice de desempenho. Algumas melhorias obtidas em outras versões são apresentadas em [2].

Um tempo de ciclo comum fixo é considerado para todas as interseções da rede com a finalidade de permitir a coordenação entre os semáforos. O coordenação é obtida através do ajuste entre os *offsets* de forma a garantir que um pelotão compacto chegue à interseção a jusante durante seu período de indicação verde [57].

O método procede de forma iterativa. Para valores conhecidos de variáveis de decisão (*splits*, *offsets*, e tempo de ciclo), o modelo de rede dinâmico calcula o índice de desempenho correspondente. O algoritmo de otimização utiliza uma heurística *hill-climb* que introduz pequenas alterações nas variáveis de decisão, requisita novo cálculo das variáveis através do modelo e mantém este procedimento até encontrar um mínimo local.

As demandas de fluxo são obtidas através de contagens veiculares realizadas manualmente nos

links de entrada. Estas contagens são descritas na forma de um histograma das chegadas de veículos que projeta o comportamento do fluxo na linha de parada, a cada ciclo. Este histograma é denominado Perfil de Fluxo Cíclico (PFC). Um exemplo deste perfil é apresentado na Figura 3.5.

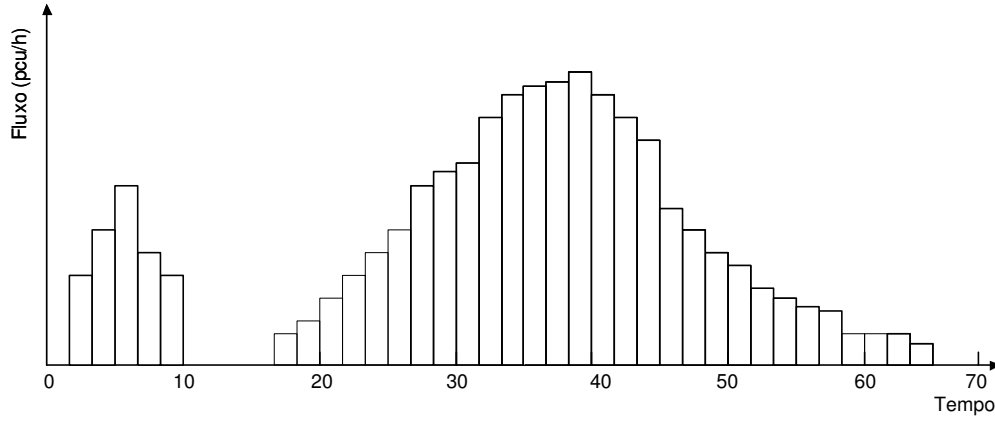


Figura 3.5: Exemplo de um Perfil de Fluxo Cíclico

Os dados numéricos são obtidos da manipulação das informações que os PFC fornecem. Assim, o número de veículos na linha de parada (m_i), durante um intervalo de tempo i é calculado pela relação:

$$m_i = \text{Max}(m_{i-1} + q_i - s_i, 0) \quad (3.6)$$

onde:

q_i - é o número de veículos que chegam na via durante um intervalo i . Este valor é dado pela soma do fluxo veicular parado durante o tempo de vermelho mais o fluxo veicular que parte durante o tempo de verde, chamado de *IN-profile* e representado pela área sob a linha escura, na Figura 3.6;

s_i - é o número máximo permitido de veículos que saem durante um intervalo i , valor dado pela soma do fluxo veicular que parte durante o tempo de verde, à razão de fluxo de saturação, mais o fluxo veicular que parte durante o tempo de verde à razão de chegada, chamado de *GO-profile* e representado pela área sob a linha escura, na Figura 3.7.

Na Figura 3.8 são representados, sob a curva escura, os fluxos de veículos que partem da linha de parada após terem sofrido atraso, este fluxo é denominado *OUT-profile*.

Outro aspecto do modelo de tráfego desenvolvido para o TRANSYT é a forma como é avaliado desempenho do sistema. é utilizada uma composição dos critérios atraso e número de paradas. O atraso é calculado da seguinte forma:

$$\rho + \Delta = \frac{T}{4} \left\{ \left[(f - F)^2 + \frac{4f}{T} \right]^{\frac{1}{2}} + (f - F) \right\} \text{pcu} - \text{horas/hora} \quad (3.7)$$

onde:

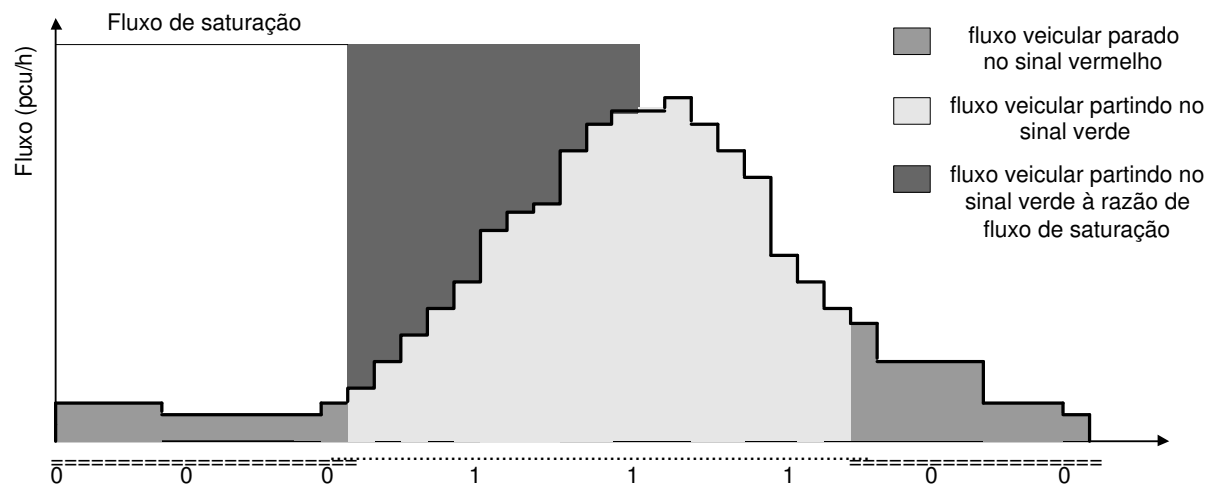


Figura 3.6: Exemplo de um PFC, *IN-profile*.

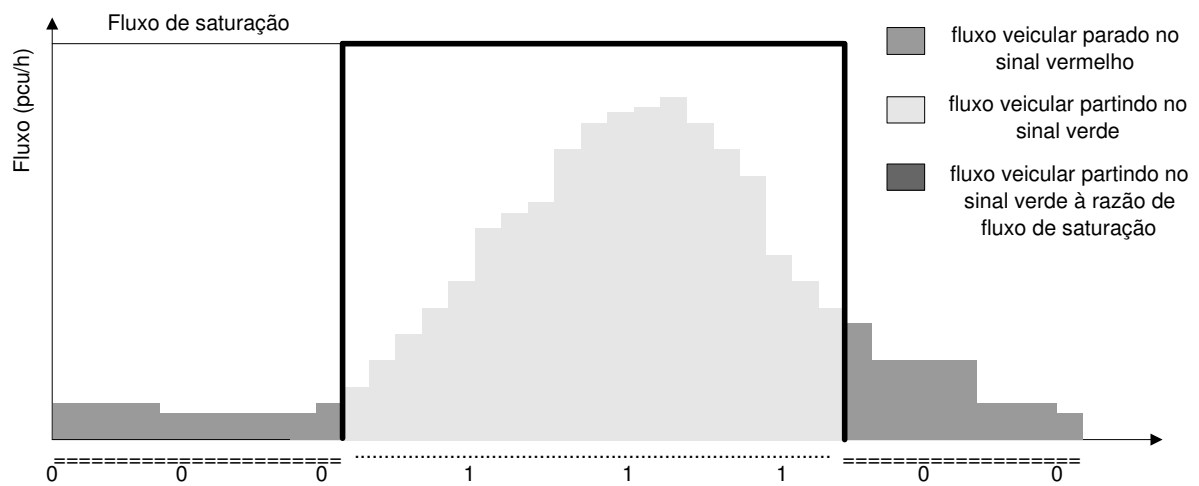


Figura 3.7: Exemplo de um PFC, *GO-profile*.

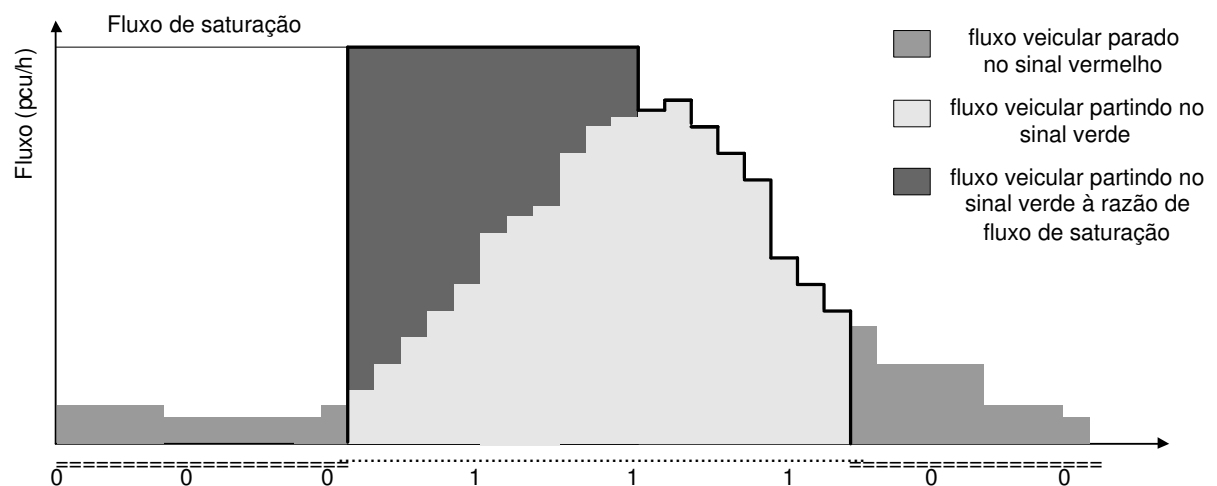


Figura 3.8: Exemplo de um PFC, *OUT-profile*.

- ρ é o atraso randômico;
- Δ é a razão de atraso quando o sistema está sobre-saturado;
- f é a razão de chegada média na via (pcu/h);
- F é o fluxo máximo que pode ser descarregado de uma via (pcu/h);
- T é a duração da condição de fluxo para cada temporização semafórica considerada (h).

A parcela uniforme do atraso é calculada de maneira similar ao primeiro termo da equação (3.2).

O critério de parada é a razão total de veículos que são forçados a parar em uma via calculado como a composição da parcela uniforme de paradas mais a parcela randômica de sobre-saturação.

Todo tráfego que sofre um atraso uniforme contribui para uma parcela do critério de parada uniforme mas o atraso no tráfego de alguns poucos segundos provoca apenas uma perda de velocidade e não paradas. Conseqüentemente, uma correção inclui uma fração de parada para pequenos atrasos, esta fração depende do tamanho do atraso.

é feita uma estimativa de paradas adicionais causadas por variações randômicas das chegadas de veículos a cada ciclo e também pela crescente saturação de filas nas vias onde as chegadas excedem a capacidade.

A estratégia TRANSYT permite a coordenação entre interseções para promover a progressão de pelotões identificados nos PFCs através do ajuste de *offsets* em seu plano de tempos fixos. A descarga de filas também é contemplada através do ajuste do *split*.

3.2.2 TRANSYT-AUT

O TRANSYT-AUT é uma modificação na estratégia anterior que busca atender a demanda do tráfego através da seleção dinâmica de planos. é utilizada a monitoração do comportamento de tráfego para definir qual melhor plano de tempos fixos a ser aplicado para a demanda detectada. Este plano é selecionado a partir de uma biblioteca de planos pré-calculados. Esta versão foi aplicada em Gothenburg, na Suécia [34].

Os princípios para obtenção de coordenação e descarga de filas são os mesmos aplicados na versão original do TRANSYT.

3.2.3 MAXBAND

A primeira versão da estratégia MAXBAND foi desenvolvido por Little, [40], ver também [41]. MAXBAND é um modelo de otimização de banda máxima que desenvolve planos de tempos semafóricos para vias arteriais. O algoritmo de otimização usado pelo MAXBAND é baseado em uma formulação de Programação Inteira Mista. Algumas características do MAXBAND são:

- a duração do ciclo é tratado como uma variável contínua sem especificação de limites;
- a velocidade dos veículos pode variar dentro de limites especificados;
- a melhor sequência de fases para cada interseção é automaticamente selecionada para um grupo específico;
- o tempo de limpeza de filas permite que o fluxo secundário acumulado durante o tempo de vermelho descarregue antes da chegada do pelotão;
- o modelo aceita comprimentos especificados pelo usuário para a banda verde em cada direção.

As entradas básicas do MAXBAND incluem a duração do ciclo, a geometria de diferentes vias, as razões de fluxo, as razões de fluxo de saturação, sequências de fases permitidas, tempos de limpeza de filas e velocidades. As saídas incluem um relatório de dados e um relatório de soluções que contém comprimentos de ciclo, comprimentos de bandas, sequências de fases selecionadas, *splits*, *offsets*, velocidade na via e tempo de viagem. Os cálculos para obtenção de tais soluções são realizados *off-line*.

O método é baseado em dados históricos e seus cálculos são realizados e tabulados previamente. Desta forma, a coordenação semaforica obtida através da otimização da banda máxima será tão boa quanto os dados históricos forem próximos da realidade. Esta otimização não garante a minimização do atraso e o número de paradas, pois não leva em conta as possíveis filas formadas durante o vermelho.

3.2.4 SCOOT

O SCOOT, *Split, Cycle and Offset Optimization Technique*, é uma ferramenta para controle de tráfego urbano desenvolvida na Inglaterra pelo *Transport and Road Research Laboratory* [36]. Ele é um sistema adaptativo que responde automaticamente às variações de demanda do tráfego. SCOOT é aplicado em mais de 170 cidades na Inglaterra e em diferentes países no mundo [44], [14]. Seu modelo e procedimento de otimização são similares aos do TRANSYT. Sendo o TRANSYT uma estratégia de controle de tempos fixos, pode-se classificá-lo como um modelo de controle em malha aberta, já o SCOOT, por ser uma estratégia de controle adaptativo, pode-se considerá-lo um modelo de controle em malha fechada, (assim como as outras estratégias que serão discutidas de controle adaptativo) [17]. O componente fundamental do TRANSYT é o PFC. No SCOOT o PFC também é aplicado, mas são utilizados detectores para medir o fluxo em tempo real. Estes detectores são localizados na entrada das vias que levam à interseção em interesse. Através dos dados do detector é obtido o perfil do pelotão de veículos que vai percorrer a via até a interseção. Este PFC é continuamente atualizado, a cada 4 segundos. Junto com o fluxo de saturação e o tempo de viagem na via, eles são utilizados para prever a fila que vai se formar na linha de parada da interseção [61].

O levantamento do PFC em tempo real pode ser analisado com auxílio da Figura 2.9 apresentada no Capítulo 2. Conhecendo-se o tamanho L da via e a velocidade de cruzeiro dos veículos, pode-se

determinar o tempo τ que os veículos levam para atingir a linha de parada, ou o último veículo parado em fila. Considerando-se um tamanho médio para os veículos, calcula-se também o tamanho da fila que se forma durante o tempo de vermelho. A descarga da fila é considerada à razão do fluxo de saturação.

Com as informações obtidas e os cálculos realizados, são feitas alterações incrementais frequentes nos valores de *split* e *offset* com o objetivo de minimizar o comprimento de fila estimado e atender ao fluxo detectado. O tempo de ciclo também é ajustado com objetivo de manter o grau de saturação em torno de 90% nas interseções. Estas alterações minimizam os transientes mas podem criar dificuldades na manutenção de uma boa coordenação [35].

Os valores das alterações incrementais são calculadas por três otimizadores:

otimizador de *split* - alguns segundos antes da mudança de fase é realizado o cálculo para três situações distintas: aumento do tempo de verde em alguns segundos, diminuição ou se mantém no valor atual. O objetivo é minimizar o grau de saturação máximo nas aproximações da interseção;

otimizador de *offset* - opera em cada interseção e em cada ciclo. A informação do PFC é usada para estimar se uma alteração no *offset* será implementada ou não para promover uma progressão do tráfego nas vias que se encontram imediatamente a montante (que descarregam fluxo na interseção, se considerado fluxo em um único sentido) ou a jusante (que recebem fluxo da interseção) da interseção em consideração. A função objetivo deste otimizador é a minimização do atraso, paradas e congestionamento;

otimizador de ciclo - o modelo SCOOT implementa uma duração de ciclo comum para cada seção (sub-rede) de semáforos com objetivo de obter coordenação entre os semáforos. O otimizador de ciclo pode variar a duração do ciclo para cada sub-rede, em incrementos de poucos segundos (4 a 8), em intervalos não menores que 2 minutos e 1/2 a 5 minutos. A duração do ciclo é otimizado com o critério de maximiza o grau de saturação para operar as interseções com um alto carregamento, em torno de 90%.

A otimização incremental dos tempos semaforicos permite um plano de coordenação que responda às novas situações de tráfego em uma série frequente, mas pequena, de incrementos. O SCOOT aplica um plano de coordenação *elástico*, ou seja, que pode ser esticado ou encolhido para satisfazer a última situação verificada pela atualização do PFC.

3.2.5 SCATS

SCATS, *Sydney Coordinated Adaptive Traffic System*, é uma estratégia que busca a otimização da operação do tráfego verificada pelos critérios de atraso e paradas de veículos pelo controle dos parâmetros ciclo, *offset* e *split*. Mesmo não sendo utilizado um modelo matemático acoplado ao otimizador, a disponibilidade de dados que descrevem o comportamento do tráfego através dos detectores são fundamentais e suficientes para a operação do algoritmo [42].

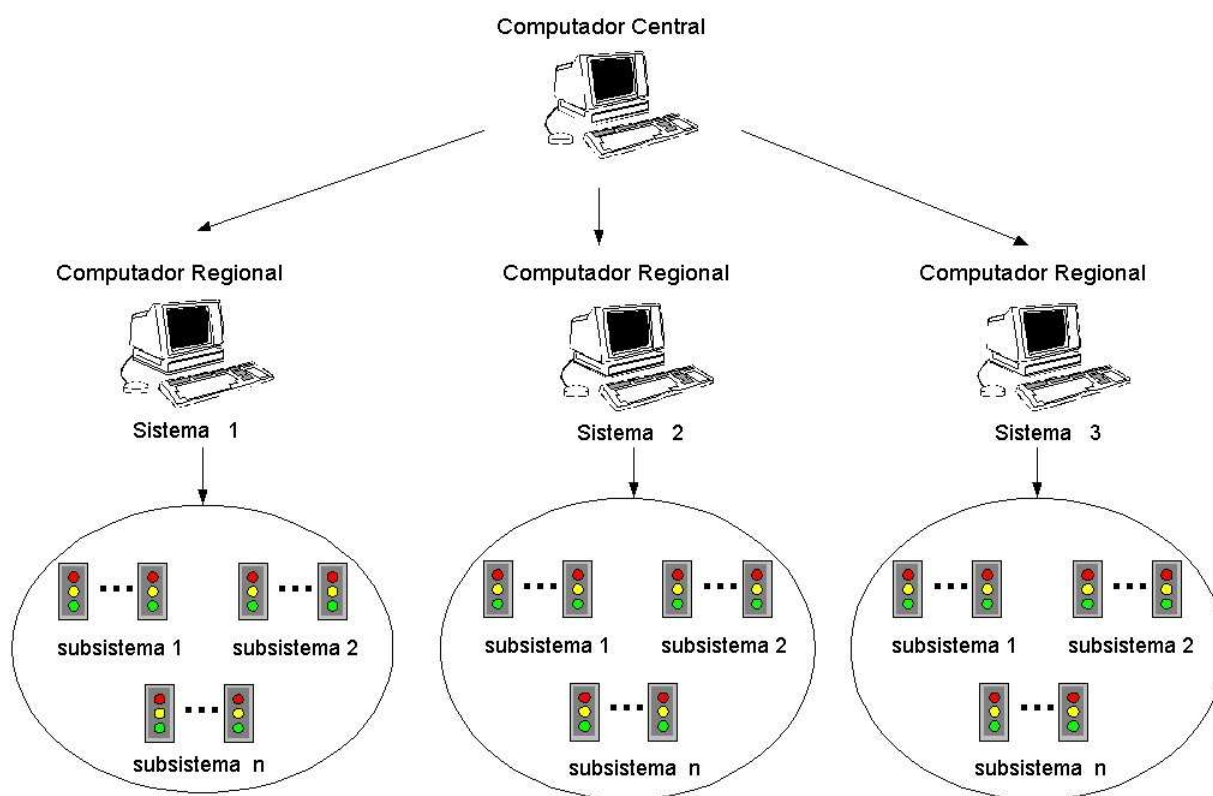


Figura 3.9: Sistema utilizado pelo SCATS.

O sistema é composto de computadores locais (regionais) e uma central. Cada computador regional mantém o controle de tráfego autônomo. O computador central é responsável pelo monitoramento do desempenho do sistema e do estado dos equipamentos. Os controladores locais são agrupados em sistemas (grupos de semáforos que não interagem com outros grupos) e subsistemas (semáforos que interagem entre si formando sistemas). Esta estrutura pode ser vista na Figura 3.9. Os subsistemas são considerados como elementos básicos de controle. Eles são compostos de uma a dez interseções que compreendem uma entidade de tráfego discreta. A estratégia de controle é selecionada pelo algoritmo, em resposta às variações detectadas de demanda e capacidade e tem como resposta os tempos de verde, *splits*, *offsets* e comprimentos de ciclos apropriados para cada subsistema e *offsets* que são aplicados entre subsistemas.

Uma biblioteca contendo quatro planos de *split* é mantida para cada interseção e a seleção do melhor plano a ser aplicado é feita em nível de subsistema, baseado nas necessidades da interseção crítica do subsistema. Normalmente é definida apenas uma interseção crítica para cada subsistema.

São mantidos também cinco planos internos de *offsets*, que determinam os *offsets* entre as interseções dos subsistemas, e cinco planos externos de *offsets* para as interseções que fazem vizinhança entre os subsistemas. Todas as interseções de um subsistema operam em uma duração de ciclo comum.

Todas as interseções e seus acessos são equipados com detectores de veículos de laço indutivo. Eles são localizados imediatamente à frente da linha de parada e têm dupla função ao proporcionar os dados de fluxo de tráfego, (i) para estratégia de controle e (ii) atuação local ou tática dos veículos.

Como estratégia de operação, os controladores locais passam para o computador regional, de cada detector definido como detector estratégico, o número de veículos contados durante o verde em cada aproximação e o tempo total em que o laço ficou desocupado durante o verde. Esta informação é utilizada para selecionar, em um ciclo, o plano de fases do *split*, o plano interno e o externo de *offset* e a duração do ciclo a ser aplicado ao subsistema para o ciclo seguinte junto com uma modificação incremental de *splits* e *offsets*.

Uma medida de tráfego básica utilizada pelo SCATS é análoga ao grau de saturação em cada aproximação da interseção. Ela é definida como a razão entre o tempo de verde efetivamente utilizado e o tempo total de verde disponível em cada aproximação. A duração de ciclo é ajustado para manter o grau de saturação em torno 0.9 na pista com o maior grau de saturação. Esta duração do ciclo é calculada para uma interseção crítica definida pelo usuário [25].

No SCATS existem quatro planos de coordenação que são tipicamente aplicados:

- baixo congestionamento;
- fluxo de tráfego balanceado;
- fluxo de tráfego desbalanceado em uma direção;
- fluxo de tráfego desbalanceado na direção oposta.

SCATS seleciona um plano de coordenação baseado nas medidas de fluxo. Para cada plano de coordenação existem dois *offsets*. Um é usualmente baseado no tempo de viagem e é usado para pequenos ciclos (e baixo congestionamento), o outro é usualmente mais curto usado para ciclos maiores (quando existe congestionamento) [25]. No entanto, estes planos de coordenação sofrem ajuste em função do grau de saturação verificado.

3.2.6 PRODYN

O PRODYN é uma estratégia de controle em tempo real, baseado em um algoritmo de Programação Dinâmica [62], associado à estratégia de horizonte deslizante para otimização dos tempos semafóricos. Os semáforos são controlados a cada período de amostragem k , de 5 segundos. Durante um dado período de amostragem, o controle a ser aplicado no período de amostragem seguinte é calculado com base nas medidas de fluxo relativas ao período de amostragem prévia. Este procedimento, conhecido por horizonte deslizante, é ilustrado na Figura 3.10.

O algoritmo constrói uma árvore de decisão com base em um modelo simplificado do tráfego baseado em equações de estado. O critério de desempenho a ser minimizado pela otimização é a soma dos atrasos sobre o horizonte mais um custo terminal que estima o atraso associado a um dado estado no final do horizonte.

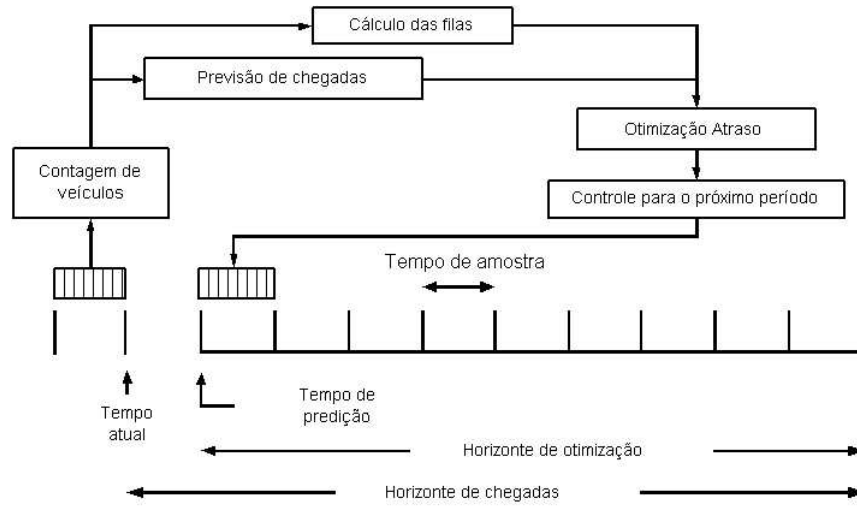


Figura 3.10: Cálculo do controle - Horizonte Deslizante (Adaptada de [22]).

O equacionamento do modelo de tráfego dinâmico, através de equações de estado, é mostrado a seguir. Inicialmente são definidas variáveis para cada interseção:

$$v_i(k+1) = u_i(k) \quad (3.8)$$

onde $v_i(k)$ é o estado da indicação semafórica vigente da interseção i durante o período k e $u_i(k)$ é a indicação semafórica definida pela ação de controle, a ser aplicado ao período subsequente $k+1$.

Como o tempo de verde é restrito a um valor máximo e mínimo de estágio, o tempo decorrido desde a última mudança de fase $w_i(k)$ é determinado pela seguinte equação:

$$w_i(k+1) = \begin{cases} w_i(k) + 1 & \text{se } u_i(k) = v_i(k) \\ 0 & \text{se } u_i(k) \neq v_i(k) \end{cases} \quad (3.9)$$

ou seja, se a ação de controle decide pela permanência da indicação semafórica vigente a variável $w_i(k+1)$ é incrementada em uma unidade de tempo de amostragem, se, pelo contrário a ação de controle determina a alteração da indicação semafórica, a variável $w_i(k+1)$ é levada a zero.

As variáveis que modelam as vias são:

- $x_l(k)$ - é a fila vertical na $l^{\text{ésima}}$ via;
- $a_{l,i}(k)$ - é o n^o de veículos que atravessam a $i^{\text{ésima}}$ seção da via l em velocidade livre;
- $m_l(k)$ - é a razão de capacidade de saída não bloqueada.

A equação de estado da fila é:

$$x_l(k+1) = \text{Max}\{0, x_l(k) + a_{l,1}(k) - d_l(k)m_l(k)\}$$

onde $d_l(k)$ é o máximo fluxo de saída teórico. Ele depende das variáveis $v_i(k)$ e u_k^i da interseção a jusante e dos parâmetros do fluxo de saturação.

A evolução da razão de capacidade de saída não bloqueada é:

$$m_l(k+1) = \begin{cases} 1 & \text{se } x_{l'}(k) + a_{l',1}(k) < s_{ll'} \\ e_l m_l(k) + f_l & \text{se } x_{l'}(k) + a_{l',1}(k) \geq s_{ll'}(k) \end{cases} \quad (3.10)$$

onde o parâmetro $s_{ll'}$ representa o fluxo na via l' que impede a saída de alguns veículos da via l . Os parâmetros e_l e f_l descrevem a velocidade e o fator de bloqueio máximo e $x_{l'}(k) + a_{l',1}(k) \geq s_{ll'}$ é para todas as vias l' que saem da interseção i ou se a via l está em vermelho para o estágio $u_i(k)$ ou para o estágio $v_i(k)$.

Os veículos que atravessam as vias em velocidade livre são dados por:

$$\begin{cases} a_{l,1}(k+1) & = a_{l,2}(k) \\ & \vdots \\ a_{l,N(l)}(k+1) & = a_{l,N(l)+1}(k) + (1-r_l)z_l(k) \\ a_{l,N(l)+1}(k+1) & = r_l z_l(k) \end{cases}$$

Os parâmetros $N(l)$ e r_l são respectivamente a parte inteira e decimal do tempo de viagem em velocidade livre expresso em termos de unidades de amostras de tempo. A variável $z_l(k)$ representa a entrada de fluxo na via l . O fluxo é medido por um detector localizado na entrada da via. Se existe uma interseção controlada imediatamente a montante, o fluxo entre as interseções é relacionado pela variável:

$$z_l(k) = \sum_{l' \in U_p(l)} p_{l',l} \text{Min}\{x_{l'}(k) + a_{l',1}(k), d_{l'}(k)m_{l'}(k)\}$$

onde $U_p(l)$ é o conjunto de vias a montante de l e $p_{l'l}$ é a razão de veículos que convergem de l' para l .

A coordenação entre interseções é realizada de forma implícita pelo PRODDYN [22]. Em nível de rede, a estrutura de controle é descentralizada. Quando o controlador de uma interseção termina sua otimização sobre o horizonte, ele simula todas as saídas da interseção relativas ao controle ótimo para todo horizonte. Tais saídas são computadas para as vias de saída e as proporções de conversão entre vias são determinadas *off-line*. Uma mensagem com 15 valores de saídas, é enviado para cada interseção controlada a jusante. Esta mensagem é utilizada pelo controlador a jusante para, no período de amostragem seguinte, calcular uma melhor predição de chegadas do que a média dos valores utilizada quando a interseção a montante é muito distante (distância maior de 200 m). Este procedimento é ilustrado pela Figura 3.11.

3.2.7 OPAC

OPAC, *Optimized Policies for Adaptive Control*, é um algoritmo de otimização em tempo real cuja arquitetura de controle é descentralizada, ou seja, ele otimiza as interseções individualmente. O

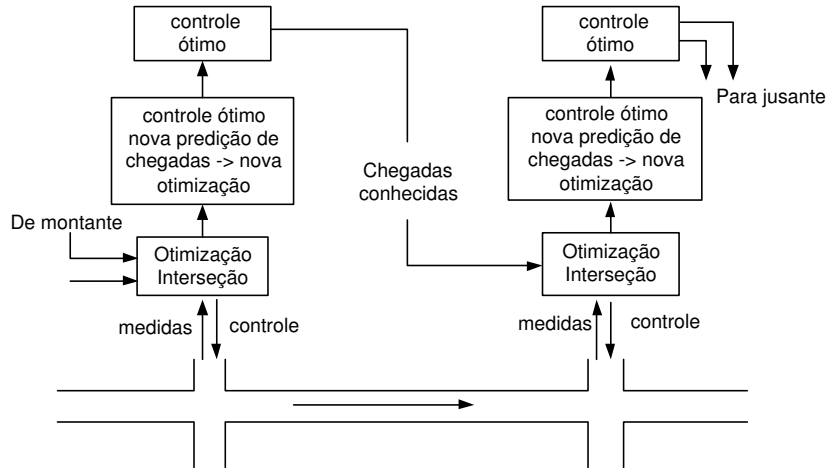


Figura 3.11: PRODYN - Coordenação implícita

algoritmo aplica Programação Dinâmica [62] na minimização do atraso total e número de paradas. Utiliza um modelo de tráfego para obter a demanda de fluxo através das medidas *on-line* e aplica estes dados na determinação da duração das fases, restritas apenas ao verde mínimo e máximo [29].

Foram desenvolvidas algumas versões até chegar na mais atualizada com opção de um módulo de coordenação. A primeira versão, designada OPAC-1, foi desenvolvida para servir de base para o desenvolvimento de estratégias futuras do próprio OPAC. é utilizado o método de Programação Dinâmica para otimização, resultando em um ótimo global para uma interseção isolada, o que requer, no entanto conhecimento de todas as chegadas durante o período de controle. Esta característica não permite sua aplicação em tempo real. Na versão 2, OPAC-2, é verificado o desempenho de cada interseção para cada período de controle através de um procedimento de otimização sequencial de busca restrita. É uma busca exaustiva de todas as possíveis combinações válidas de mudanças de fase em um período de controle para determinar um grupo de resultados ótimos. A esta versão foi incorporado a técnica de horizonte deslizante, OPAC-3, tornando possível sua utilização em tempo real. A versão 4, OPAC-4, inclui o módulo de coordenação [30].

O funcionamento do algoritmo OPAC pode ser brevemente descrito como um processo de otimização dividido em tempos de controle sequenciais de 50 a 100 segundos. Durante cada tempo de controle pode existir de uma a três mudanças de fase. A função objetivo é avaliada para todas as sequências viáveis.

Para cada etapa i tem-se o vetor de entradas de estado I_i , o vetor de chegadas de veículos A_i , o vetor de saídas de veículos O_i , a variável de decisão x_i ($x_i = 1$ se a temporização sofrer mudança na etapa ou $x_i = 0$ se permanecer o estado atual), o custo r_i (o atraso total) e as transformadas:

$$O_i = T_i(I_i, A_i, x_i)$$

$$r_i = R_i(I_i, A_i, x_i)$$

onde T_i é a relação entre as variáveis de entrada e as de saída baseada no processo de descarga das filas na interseção.

A função de otimização recursiva é dada por:

$$f_i^*(I_i) = \text{Min}_{x_i} [R_i(I_i, A_i, x_i) + f_{i+1}^*(I_{i+1})]$$

onde $I_{i+1} = G(I_i, A_i, x_i)$.

Para uma dada sequência de etapas n é definida uma função desempenho para cada aproximação, que calcula o atraso total durante a etapa:

$$\Phi_n(t_1, t_2, t_3) = \sum_{i=1}^k [Q_0 + A_i - D_i]$$

onde Q_0 é a fila inicial, A_i são as chegadas durante o intervalo i , D_i são as partidas durante o intervalo i e (t_1, t_2, t_3) são os possíveis chaveamentos durante cada etapa. Note que o número de chaveamentos durante o horizonte não pode ser menor do que um e nem maior do que três [51].

A técnica de horizonte deslizante tem algumas características de aplicação para o algoritmo OPAC: o tamanho do horizonte é dividido em k intervalos e subdivididos em r e $k - r$ intervalos. Para os detectores a montante pode-se obter os dados de chegadas atuais para um período de r intervalos, que são o *início* do estágio. Para os $(k - r)$ intervalos restantes, o *restante* do estágio, são obtidos os dados através do modelo de tráfego. É calculada a política ótima para o estágio inteiro, mas implementado somente para o *início*. Então é rolado o horizonte para frente em r unidades, obtém-se novos dados de fluxo para o estágio (*início e restante*) e repete-se o processo [26]. Um estágio n pode ser ilustrado na Figura 3.12.

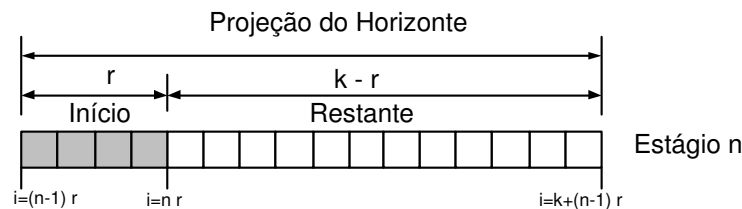


Figura 3.12: OPAC - Horizonte Deslizante (adaptado de [29])

A versão mais recente do algoritmo, OPAC-4, mantém as características já descritas, com o diferencial de oferecer um módulo de coordenação [28]. Sua arquitetura passa a ser multi-nível com um Ciclo Fixo Virtual (VFC). Este ciclo é virtual e calculado *on-line*. Ele é imposto como restrição para todos os controladores locais, permitindo flexibilidade para o começo e término do ciclo em cada interseção. Também provê uma tolerância máxima para os controladores individuais respeitando a demanda local e mantém a capacidade de coordenação nas interseções adjacentes promovendo a progressão de pelotões.

O VFC é determinado com os dados da interseção dominante. A identificação desta interseção é realizada em tempo real. O número de vezes que o cálculo *on-line* do VFC feito é especificado pelo usuário.

Os níveis de controle são três e podem ser vistos na Figura 3.13. No primeiro o controlador local

implementa o procedimento de horizonte deslizante sujeito à restrição do ciclo fixo virtual (VFC) em comunicação com o terceiro nível. O nível dois, nível de coordenação, otimiza os *offsets* de cada interseção (um por ciclo). O nível 3, nível de sincronização, calcula para a rede toda o VFC (e períodos especificados pelo usuário).

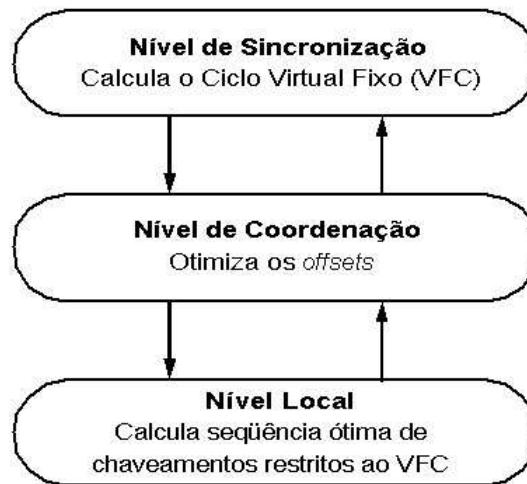


Figura 3.13: Níveis de Coordenação - OPAC

3.2.8 RHODES

O sistema RHODES [45] utiliza uma arquitetura de controle que:

- decompõe o problema de controle de tráfego em vários sub-problemas inter-conectados hierarquicamente;
- prevê os fluxos de tráfego em níveis de resolução apropriados para permitir um controle pró-ativo;
- permite vários módulos de otimização para resolver hierarquicamente os sub-problemas;
- utiliza uma estrutura de dados e comunicação/computação que permite uma solução rápida dos sub-problemas antes que a decisão possa ser enviada a campo.

A estrutura do sistema RHODES é mostrada na Figura 3.14. Ela é dividida em três níveis: o controle de rede, controle de fluxo e controle de interseções. Para o controle de interseções o algoritmo utiliza a Programação Dinâmica com horizonte deslizante. São elaborados planos de tempos em termos de duração de fase para uma dada seqüência de fases. Os dados de entrada são os fluxos medidos em tempo real por detectores e controladores de fluxo otimizados. O critério de desempenho aplicado é a combinação de atraso com número de paradas.

A estimação das filas é dada por:

$$q(t_1) = q(t_0) + a(t_1, t_0) - d(t_1, t_0)$$

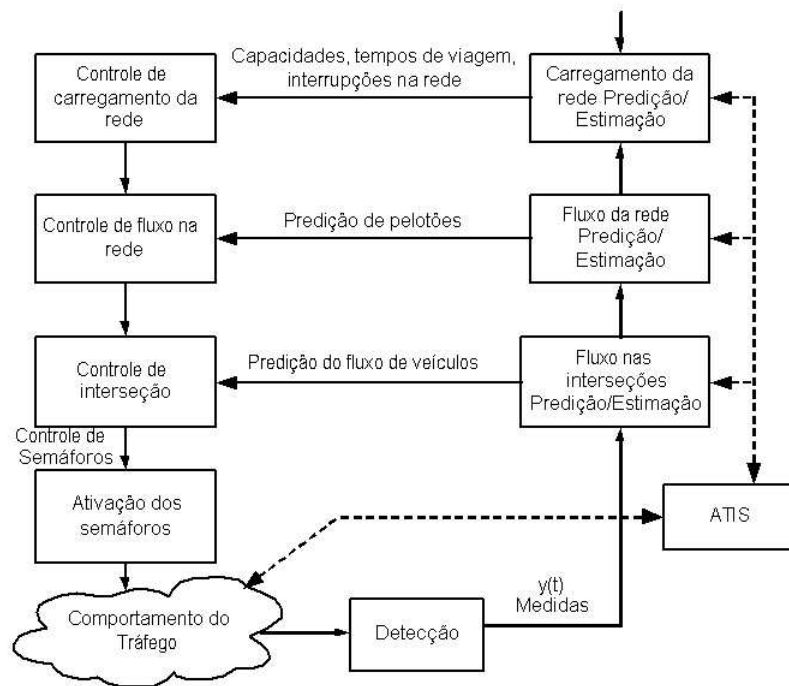


Figura 3.14: RHODES - Arquitetura hierárquica

onde $q(t)$ é a variável que determina o tamanho da fila $a(t_1, t_0)$ é o número da predição de chegadas entre o intervalo (t_1, t_0) , e $d(t_1, t_0)$ é o número de predição de partidas (usando uma da razão de descarga de filas conhecida).

A coordenação entre interseções é realizada em nível de rede pelo controle de fluxo de rede, baseado no modelo chamado REALBAND [18] que otimiza o movimento dos pelotões observados na subrede. Ao considerar pelotões, ao invés de tempos de verde, o REALBAND otimiza em relação ao fluxo, e não às indicações semaforicas, como é o caso do MAXBAND. O algoritmo tenta formar bandas de progressão baseada nas observações *on-line* realizadas na rede. Se a minimização do critério de desempenho número de paradas for obtido o algoritmo tenta formar bandas de progressão pela observação atualizada dos pelotões na rede.

3.2.9 ALLONS-D

ALLONS-D, *Adaptive Limited Lookahead Optimization of Network Signals - Decentralized*, é um esquema de controle de tráfego em tempo real. Seu algoritmo de otimização é baseado em uma busca em árvore primeiro em profundidade depois *branch and bound* com ordenação de ramos [62]. A busca é restrita apenas pelo máximo e mínimo tempo de verde por fase. O conceito de horizonte deslizante é utilizado para o cálculo e implementação dos planos semaforicos otimizados. No controlador é agregado um otimizador que utiliza informações sobre o tamanho das filas e das chegadas futuras em suas aproximações para definição da melhor temporização a ser aplicada. Estas informações são obtidas através de detectores e de sofisticado esquema de câmeras como AUTOSCOPE [52].

O modelo de tráfego utilizado define um nó da árvore de decisão como o início e o fim de cada

intervalo de tempo. Estes nós representam o estado da interseção. Cada estado é denotado por um vetor \bar{X} descrito por:

$$\bar{X}(d) = \begin{bmatrix} x_1(d) \\ \bar{x}_2(d) \\ x_3(d) \\ x_4(d) \end{bmatrix} = \begin{bmatrix} t \\ \bar{q}(t) \\ p(t) \\ g(t) \end{bmatrix} \quad (3.11)$$

onde t é o tempo absoluto em cada nó da árvore; $\bar{q}(t)$ é o vetor do tamanho das filas em todas as aproximações, $\bar{q}(t) = [q^1 \dots q^{l_{max}}]$; l_{max} é o número máximo de vias que chegam na interseção; $p(t)$ é a fase verde vigente; $g(t)$ é o tempo que a fase verde permanece ativa e d representa a busca em profundidade, sendo $d = 0$ a raiz da árvore e seus incrementos $d = d + 1$, realizados a cada nó.

A busca na árvore é realizada e, em cada nó, calculada a atualização dos estados caracterizados por:

$$\bar{X}(d+1) = f(\bar{X}(d), u(d), A(\cdot))$$

onde $u(\cdot)$ é a variável de controle, com $u(\cdot) \neq 0$ ocorre mudança de fase ou $u(\cdot) = 0$ quando a fase é mantida; $A(\cdot)$ é o vetor que representa as chegadas de veículos dentro de um certo intervalo de tempo. Este vetor consiste de chegadas em cada aproximação da interseção. Somente chegadas dentro de um intervalo de tempo específico são consideradas, este intervalo de tempo é definido pelo controle $u(\cdot)$, pelo estado existente $X(\cdot)$ e pelo tempo t . As equações de atualização dos estados são:

$$x_1(d+1) = x_1(d) + t_{step} \quad (3.12)$$

$$\bar{x}_2(d+1) = \bar{x}_2(d) - D(\bar{x}_2(d), u(d), A(t, t_{step})) \quad (3.13)$$

$$x_3(d+1) = x_3(d) + u(d) \quad (3.14)$$

$$x_4(d+1) = \begin{cases} t_{MG} & \text{se } u(\cdot) \neq 0 \\ x_4(d) + t_{step} & \text{se } u(\cdot) = 0 \end{cases} \quad (3.15)$$

O número de partidas $D(\cdot)$ é função da variável de controle, do tamanho de fila no intervalo anterior e das chegadas durante o intervalo atual; t_{MG} é uma constante que representa o tempo mínimo contínuo para uma fase tornar-se verde; t_{step} é o período de tempo para uma simples decisão de controle ($t_{step} = t_{end} - t_{start}$) e atualizado pelo controle:

$$t_{step} = \begin{cases} \Delta & \text{se } u(\cdot) = 0 \\ t_{MG} + t_{YR} & \text{se } u(\cdot) \neq 0 \end{cases} \quad (3.16)$$

onde Δ é o período de tempo, tipicamente 5 a 15 segundos, quando ocorre a coleta de dados por imagem; t_{YR} é uma constante que representa o tempo perdido, incluindo o tempo de amarelo.

A função objetivo a ser otimizada é o atraso acumulado de todos os veículos que passam em uma interseção. No cálculo do atraso considera-se a soma do atraso no tempo multiplicado por um peso associado a cada veículo. Isto permite incrementar prioridade a veículos de certo tipo ou ocupação por crescimento no atraso potencial que cada veículo contribui ao diminuir sua velocidade ou parar.

Cada veículo na intersecção pode contribuir para o atraso quando a indicação está vermelha, verde

ou amarela. São cinco as situações de atraso e cinco diferentes equações [51] para representá-las:

- um veículo que espera durante a indicação vermelha;
- um veículo que está presente na fila quando a indicação torna-se verde mas não pode partir em função dos veículos que estão a sua frente;
- um veículo que está presente na fila quando a indicação torna-se verde e parte após os veículos a sua frente partirem;
- um veículo que chega durante um período de verde mas não pode partir em função dos veículos parados a sua frente;
- um veículo que chega durante o período de verde e parte depois dos veículos que estão parados a sua frente.

As primeiras versões do ALLONS-D apresentam um esquema de coordenação implícita associada às interseções que ocorre sem comandos explícitos de coordenação. Isto ocorre devido a incorporação de dados da interseção a montante, das predições de chegadas e da implementação do horizonte deslizante.

Além da coordenação implícita foi implementado posteriormente um esquema de coordenação explícita. é considerado um multiplicador parcial (um peso) a cada interseção. Este peso é adicionado à função objetivo e proposto como variável de controle. Uma definição apropriada dos pesos define uma boa coordenação. Os pesos na função objetivo funcionam de forma que o algoritmo de otimização dá uma maior consideração a fase do tráfego de maior peso [50].

Este esquema de coordenação é dividido em três níveis: o nível regional, onde ocorrem as medidas e predições, além da determinação da rede sob controle; o nível de rede, onde são determinadas as características das vias e conseqüentemente seus respectivos pesos, estes pesos são definidos a cada período de tempo; e os níveis locais, que aplicam o controle definido na otimização.

3.2.10 CRONOS

CRONOS, *ContROl of Networks by Optimization of Switchovers* [11] é um algoritmo em tempo real que inclui um método de otimização polinomial com complexidade crescente facilitando simultaneamente o controle de várias interseções e o tratamento de situações de sobre-saturação. O método de otimização é baseado no *BOX's algorithm* [38].

Suas características quanto ao critério de desempenho, obtenção do conhecimento do número de chegadas de veículos através de um modelo de tráfego são similares aos outros algoritmos já citados. Duas são as principais diferenças: a fila é modelada de forma horizontal, permitindo assim a avaliação de ocupação espacial dos veículos nas vias e a inclusão das informações das vias que partem da interseção em estudo ao modelo.

A estratégia CRONOS utiliza sensores de vídeo para obter as medidas de tráfego em tempo-real, as quais consistem na medição direta do comprimento das filas.

3.2.11 TUC

A metodologia aplicada na estratégia de controle semafórico TUC (*Traffic-responsive Urban Control*) trata do problema de desempenho de um sistema de tráfego quando este se encontra em condições de tráfego sobre-saturado. Estratégias como SCOOT e SCATS, tratam de situações de sobre-saturação quando estas são de curta duração [56], [43].

O problema de tráfego urbano foi modelado como um problema de otimização baseado no modelo matemático apresentado em [31]. O objetivo do controle de TUC é a minimização e equilíbrio da quantidade de veículos nas vias de uma dada rede controlada. A rede é representada por um grafo direcionado com vias z , com $z \in Z$ e interseções j , com $j \in J$. Para cada interseção controlada j , I_j e O_j representam os grupos de entradas e saídas da interseção j , respectivamente. Assume-se que o *offset*, o ciclo C_j , e o tempo perdido total L_j da interseção j são fixos; por simplificação, assume-se $C_j = C$ para todas as interseções. O controle em uma interseção j é baseado em um número fixo de estágios que pertence ao grupo F_j , enquanto v_z é o grupo de estágios cujas vias z tem prioridade de passagem. Finalmente, o fluxo de saturação S_z , $z \in I_j$, e as razões de conversão $t_{z,w}$, $z \in I_j$, $w \in O_j$, são conhecidas e fixas.

Por definição, para a interseção j , a seguinte restrição se aplica:

$$\sum_{i \in F_j} g_{j,i} + L_j = C \quad (3.17)$$

onde $g_{j,i}$ é o tempo de verde do estágio i da interseção j . A variável $g_{i,j}$ é restrita pelo verde máximo, e mínimo, $[g_{j,i,min}, g_{j,i,max}]$.

Considerando uma via z conectando duas interseções M e N tal que $z \in O_M$ e $z \in I_N$, representada na Figura 3.15. A dinâmica da via z é dada por:

$$x_z(k+1) = x_z(k) + T[q_z(k) - s_z(k) + d_z(k) - u_z(k)] \quad (3.18)$$

onde x_z é o número de veículos na via z ; q_z e u_z são os fluxos de entrada e saída, respectivamente, da via z no período $[kT, (k+1)T]$ com T sendo o intervalo de controle e $k = 1, 2, \dots$ o índice que representa os intervalos discretos do tempo; d_z e s_z são a entrada e saída de fluxo que pode ocorrer no meio da via.

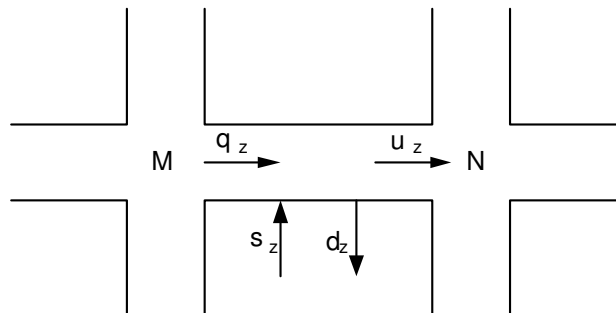


Figura 3.15: Via de uma rede urbana

O fluxo de saída é dado por:

$$s_z(k) = t_{z,0}q_z(k) \quad (3.19)$$

onde $t_{z,0}$ é considerada fixa e conhecida.

O fluxo de entrada na via z é dado por:

$$q_z(k) = \sum_{w \in I_M} t_{w,z}u_w(k) \quad (3.20)$$

O intervalo de controle T é escolhido não menor que o ciclo C , um valor médio é obtido por $u_z(k) = \frac{S_z G_z(k)}{C}$, onde G_z é o tempo de verde efetivo na via z . Assumindo disponíveis os tempos de verde nominais $g_{j,i}^N$ e a demanda nominal d_z^N , tem-se da equação (3.18):

$$(1 - t_{z,0})q_z^N + d_z^N - u_z^N = 0 \quad (3.21)$$

Introduzindo todas as considerações na equação (3.18), tem-se:

$$x_z(k+1) = x_z(k) + T \left[(1 - t_{z,0}) \sum_{w \in I_M} \frac{t_{w,z} S_w \left(\sum_{i \in v_w} \Delta g_{M,i}(k) \right)}{C} + \Delta d_z(k) - \frac{S_z \left(\sum_{i \in v_z} \Delta g_{N,i}(k) \right)}{C} \right] \quad (3.22)$$

onde $\Delta g_{j,i} = g_{j,i} - g_{j,i}^N$ e $\Delta d_z = d_z - d_z^N$.

Aplicando a equação (3.22) a uma rede arbitrária, as equações de estados descrevem a evolução do sistema no tempo:

$$\mathbf{x}(k+1) = \mathbf{A}\mathbf{x}(k) + \mathbf{B}\Delta\mathbf{g}(k) + \mathbf{T}\Delta\mathbf{d}(k) \quad (3.23)$$

onde \mathbf{x} é o vetor de estados do número de veículos na via $z \in Z$; $\Delta\mathbf{g}$ é o vetor de estados para $\forall i \in F_j, \forall j \in J$; $\Delta\mathbf{d}$ é o vetor $\Delta d_z = d_z - d_z^N$; e $\mathbf{A} = \mathbf{I}$, \mathbf{B} e \mathbf{T} são os estados, entradas e matrizes de distúrbios, respectivamente. A matriz \mathbf{B} reflete a topologia da rede.

Para aplicação da metodologia LQ (Linear quadrática) [62], para controle realimentado, considere-se $\Delta\mathbf{d}(k) = 0$ levando a equação (3.23) a:

$$\mathbf{x}(k+1) = \mathbf{x}(k) + \mathbf{B}\Delta\mathbf{g}(k). \quad (3.24)$$

Isto é, o controle não é projetado para rejeitar explicitamente as perturbações Δd ; entretanto, por se tratar de um sistema realimentado, a rejeição ocorre ao longo do tempo.

A metodologia de Otimização LQ deve ser vista como caminho para obter eficiência com a matriz ganho melhor do que com tentativas de otimização de critérios físicos sujeitos à acuracidade do modelo de equações e restrições.

Para minimizar o risco de sobre-saturação e bloqueios para as filas, deve-se tentar minimizar e balancear a ocupação relativa das vias $x_z/x_{z,max}$, onde $x_{z,max}$ é a capacidade de uma via $z \in Z$,

medida em veículos. O critério quadrático que considera o objetivo deste controle é:

$$\mathfrak{J} = \frac{1}{2} \sum_{k=0}^{\infty} (\|\mathbf{x}(k)\|_{\mathbf{Q}}^2 + \|\Delta \mathbf{g}(k)\|_{\mathbf{R}}^2) \quad (3.25)$$

onde \mathbf{Q} e \mathbf{R} são matrizes de peso, diagonais, definidas não negativas. Os elementos da diagonal de \mathbf{Q} são o inverso da capacidade máxima $x_{z,max}$ e \mathbf{R} pode ser escolhida livremente.

A minimização do critério de desempenho da equação (3.25) sujeito a equação (3.24) leva a lei LQ de controle realimentado:

$$\mathbf{g}(k) = \mathbf{g}^N - \mathbf{L}\mathbf{x}(k) \quad (3.26)$$

onde \mathbf{L} é a matriz de controle resultante que depende das matrizes \mathbf{B} , \mathbf{Q} e \mathbf{R} , que são encontradas por simulação têm grande sensibilidade com respeito às variações dos parâmetros de tráfego [19].

A metodologia LQ não leva em consideração a existência de restrições ao controle, as restrições existentes são impostas depois da aplicação da equação (3.26). Ao final, o problema de otimização é resolvido em tempo real para cada interseção j , para tempos de verde $G_{j,i}$ específicos:

$$\begin{aligned} & \text{Min}_{G_{j,i}} \sum_{i \in F_j} (g_{j,i} - G_{j,i})^2 \\ & \text{sujeito a } \sum_{i \in F_j} G_{j,i} + L_j = C \\ & G_{j,i} \in [g_{j,i,min}, g_{j,i,max}] \forall i \in F_j \end{aligned}$$

O regulador multivariável dado em (3.26) é reativo e responde indiretamente a distúrbios desconhecidos, não sendo necessárias previsões futuras das condições de tráfego [20].

A lei de controle dada na equação (3.26) requer valores disponível de \mathbf{g}^N . Alternativamente esta lei pode ser reescrita da seguinte forma:

$$\mathbf{g}(k) = \mathbf{g}(k-1) - \mathbf{L}[\mathbf{x}(k) - \mathbf{x}(k-1)] \quad (3.27)$$

Esta nova lei de controle, equação (3.27) é a formulação do problema de tráfego como um problema de controle ótimo linear-quadrático-integral (LQI).

Ao adotar o período de amostragem igual ao ciclo, a estratégia TUC perde a propriedade de reagir a eventos locais e temporários em interseções específicas. Por outro lado, ganha-se do ponto de vista da gerência global da capacidade da rede, pois a matriz \mathbf{L} reflete a influência de interseções distantes no cálculo do controle local.

3.3 Classificação

Nas seções seguintes serão discutidas características comuns às estratégias citadas, com o objetivo de classificá-las tentando assim facilitar sua compreensão.

3.3.1 Arquitetura e Estratégia de Otimização

Interseções semaforizadas em áreas urbanas podem ser tratadas como interseções isoladas ou como parte de um sistema coordenado. O tratamento como interseções isoladas ocorre quando a distância entre as interseções é grande o suficiente para que a influência entre elas possa ser desprezada. Quando a distância entre as interseções não é desprezível, um conjunto delas é tratado como uma área de controle frequentemente coordenada por uma Central de Controle de Tráfego (CCT). Esta CCT pode aplicar estratégias de controle que buscam otimizar o fluxo entre as interseções de forma coordenada. O modelo de tráfego pode ter uma arquitetura centralizada (global), ou seja, quando considera a influência das ações de controle mútuas entre interseções vizinhas e a dinâmica do fluxo do tráfego, permitindo a otimização do controle para a rede de tráfego como um todo. Este tipo de otimização pode tornar-se inviável para aplicações práticas em função do custo computacional crescente em função da dimensão da rede. Por outro lado, a arquitetura do modelo de tráfego pode ser descentralizada, ou seja, cada interseção é tratada individualmente sendo a otimização dos tempos semafóricos realizada localmente em cada interseção.

A grande maioria das estratégias de controle possuem uma arquitetura descentralizada, como o TRANSYT, SCOOT, SCATS, PROLYN, OPAC, RHODES, ALLONS-D e o CRONOS. Os algoritmos de controle que foram desenvolvidos sobre esta arquitetura utilizam-se de algumas técnicas que provocam implicitamente a troca de informações entre as interseções. Uma delas é o horizonte deslizante utilizado para projetar o comportamento do tráfego em um horizonte futuro finito para diferentes situações de controle. Entre duas amostras de tempo é realizada esta projeção, encontrada a melhor opção dentre as testadas e o resultado aplicado. Durante a projeção do horizonte são utilizadas informações das chegadas de veículos recebidas de outras interseções e considerado também o controle existente nas outras interseções.

Já uma arquitetura centralizada tem seu representante na estratégia TUC [49].

As características do problema de controle de tráfego restringem os métodos de otimização que podem ser aplicados. O método mais utilizado, dependendo do modelo utilizado é a Programação Dinâmica, utilizada pelo PROLYN, OPAC, RHODES e ALLONS-D em sua primeira versão. Nas versões posteriores ALLONS-D aplica um algoritmo baseado em uma busca em árvore, primeiro em profundidade e depois *branch and bound* com ordenação de ramos. A técnica *Hill-Climbing* é utilizada pelo TRANSYT e pelo SCOOT. O CRONOS aplica um método polinomial conhecido por *BOX's Algorithm* [11].

3.3.2 Modelo de Fila e Detecção

O modelo de fila mais aplicado é o vertical, onde considera-se que a fila forma-se antes da linha de parada e neste ponto os veículos acumulam-se verticalmente, não havendo um limite para a quantidade de veículos em fila. Dentre todas as estratégias citadas apenas CRONOS [11] não utiliza o modelo de fila vertical.

Outro modelo de fila é o horizontal, onde é considerado um tamanho médio de veículos e o espaço que eles ocupam na via horizontalmente. Este modelo possibilita a verificação do espaço que uma fila ocupa e conseqüentemente se esta fila atinge a capacidade total da via. No sistema CRONOS a fila é observada diretamente através da captura de imagens via câmeras.

A utilização de detectores permite adaptação do controle às condições vigentes de tráfego, através da contagem de veículos. Estes detectores podem ser de laço indutivo ou através de processamento de imagens por câmeras. Contagens manuais são utilizadas pelos sistemas de tempo fixo para elaboração de planos de tempos semafóricos, pratica aplicada no TRANSYT. Já os detectores, cuja localização auxilia na verificação e predição da formação de fila, podem ter diferentes funções quanto à sua localização:

- *Saídas de todas as aproximações da interseção em análise* - informa a quantidade de veículos que sai de cada aproximação e também, o tempo total em que o laço indutivo fica desocupado durante o tempo de verde. Esta localização de detectores é utilizada pelo SCATS e pelo RHODES. A estratégia SCATS utiliza a informação de ocupação do detector para o cálculo do grau de saturação. Esta alocação de detectores pode ser verificada na Figura 3.16 (a).
- *Início da via logo após a linha de parada da via a montante da interseção* - é medido o fluxo que entra na via, considerando-se que não ocorrem paradas nem saídas durante o percurso; calcula-se os tempos semafóricos para atender o fluxo medido. Este é o caso aplicado no SCOOT e mostrado na Figura 3.16 (b);
- *Saída da aproximação a montante da interseção e outro a 50 m antes da linha de parada* - o fluxo medido na saída da interseção a montante é usado na predição de chegadas dos veículos na linha de parada. O detector localizado a 50 m acima da linha de parada é utilizado auxiliar no cálculo de estimação de filas. Este modelo de detecção é utilizado no PROLYN e mostrado na Figura 3.16 (c).

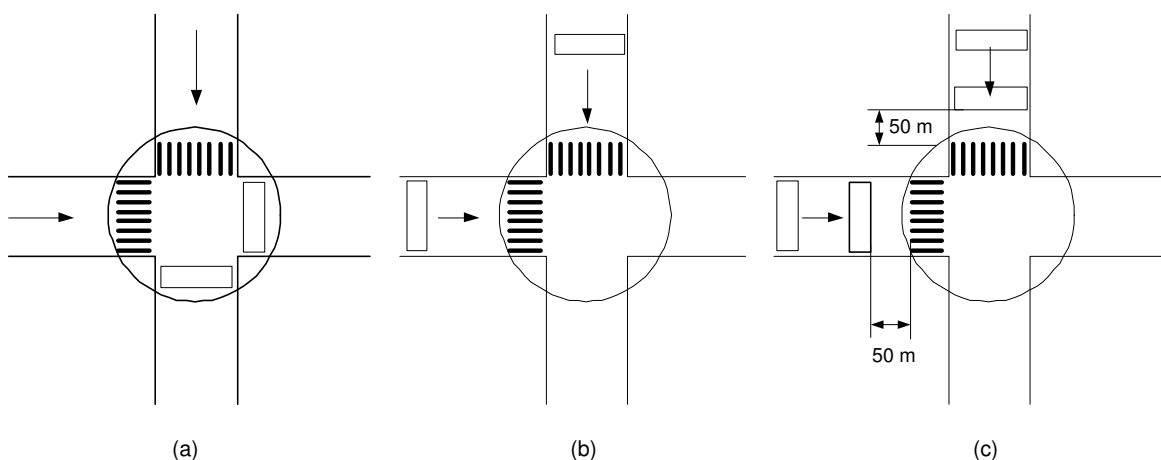


Figura 3.16: Localização de detectores

As medidas realizadas através de sensores de vídeo permitem a verificação direta do número de veículos ou a fila em uma via. Também a ocorrência de eventos especiais como acidentes e quebra

de veículos que reduzem a capacidade momentânea da via. Este tipo de medida é utilizada pelas estratégias ALLONS-D e CRONOS.

3.3.3 Critério de Desempenho

O melhor ajuste dos tempos semaforicos é obtido após avaliação de um critério de desempenho. Estes critérios já foram discutidos na seção 2.8.

As estratégias como o PROLYN, ALLONS-D aplicam apenas o atraso como critério de desempenho.

Os algoritmos restantes aplicam a combinação entre o atraso e o número de paradas.

Na tabela 3.1 será apresentado um resumo das características discutidas.

Tabela 3.1: Classificação dos Modelos de Tráfego e Estratégias de Controle.

Características	Estratégias de Controle								
	Transyt	Scoot	Scats	Prodyn	Opac	Rhodes	Allons-d	Cronos	TUC
Arquitetura	D	D	D	D	D	D	D	D	C
Otimização	H-C	H-C	He	PD	PD	PD	PD, dfs-bb	Box's	LQ
Modelo de Fila	V	V	V	V	V	V	V	H	V
Desempenho	A-P	A-P	A-P	A	A-P	A-P	A	A-P	A-P

legenda: D - Descentralizado, C - Centralizado, H-C - Hill-Climbing, PD - Programação Dinâmica, dfs-bb - busca em profundidade depois branch and bound, V - Vertical, H - Horizontal, A-P - Atraso e Parada, A - Atraso, He - Heurística, LQ - Linear Quadrática.

3.4 Conclusões

Foi apresentado um panorama das estratégias de controle de tráfego mais conhecidas e relevantes para o problema de desempenho de sistemas viários. Foi apresentado também como cada estratégia trata os problemas de coordenação e descarga de filas. Para auxiliar na compreensão das estratégias de controle existentes foram apresentadas algumas características de forma a obter uma classificação das estratégias apresentadas.

As estratégias apresentadas foram testadas e algumas aplicadas em seus países de origem. Uma grande limitação na utilização destas estratégias, como pacotes comerciais a serem adquiridos e aplicados em sistemas de tráfego urbano brasileiro é que cada uma delas têm especificidades relativas aos sistemas de tráfego de seus países de origem. Além disso, o custo destas estratégias é elevado na aquisição, implantação e manutenção. Desta forma o algoritmo proposto neste trabalho e apresentado no próximo capítulo tem características similares às estratégias apresentadas quanto a arquitetura, estratégia de otimização, modelo de fila e detecção, e ainda critério de desempenho. As simplificações

e diferenças se devem às características dos nossos sistemas de tráfego e da busca de um algoritmo para implementação em baixo custo.

Capítulo 4

Modelo de Simulação de Tráfego e Metodologia de Solução

4.1 Introdução

No Capítulo 3 foram apresentados alguns modelos de tráfego existentes na literatura, inclusive o modelo aplicado na estratégia de controle PRODYN, que foi a base para o desenvolvimento do modelo de simulação de tráfego aplicado neste trabalho.

Este capítulo apresenta o desenvolvimento do modelo de tráfego, necessário para simulação do tráfego urbano e suas características. Será apresentado também um método de solução para o problema de minimização de atraso através de um algoritmo de controle em tempo real com uma heurística de busca em profundidade.

Alguns resultados obtidos através do algoritmo de controle serão apresentados e comparados com a estratégia de tempos fixos TRANSYT. A estratégia TRANSYT é como citado em [49]: 'É a estratégia de controle semafórico mais conhecida e frequentemente aplicada, e também utilizada como um método de referência para testes de estratégia em tempo-real.' Outra razão para utilização da estratégia TRANSYT como padrão de comparação foi a aquisição do *software* TRANSYT/10 pela Universidade Federal de Santa Catarina.

4.1.1 Modelo de Tráfego Desenvolvido para Este Trabalho

Apresenta-se o modelo de tráfego utilizado neste trabalho, para simulação das condições de tráfego, não saturadas, necessárias para o desenvolvimento do algoritmo de controle. Este modelo foi desenvolvido com base no modelo de tráfego do PRODYN [21], sendo uma simplificação deste, pois, não considera aspectos como bloqueios por conversão à esquerda em vias.

No Brasil não são comuns planos semafóricos que incluam conversões à esquerda permitidas e não protegidas. Situações aqui encontradas podem ser caracterizadas por:

1. não permitidas;
2. permitidas e protegidas: é definido um estágio onde os veículos fazem a conversão com os movimentos conflitantes recebendo indicação vermelha;
3. permitidas e não protegidas: o veículo pode fazer conversão à esquerda buscando um *gap* no fluxo conflitante.

São considerados instantes de tempo t , (tempo de amostra), espaçados de um período constante t_s , no caso 4 segundos, ou seja: $t_s = t_{j+1} - t_j$. O fluxo de veículos em uma rede é modelado através de equações de estado. Há um grupo de equações de estados para interseções que inicialmente descrevem a evolução da indicação semafórica:

$$v_i(t+1) = u_i(t), \quad i = 1, 2, \dots, I \quad (4.1)$$

onde $v_i(t) \in \{0, 1\}$ é o estágio vigente na interseção i , durante o intervalo de tempo (t_j, t_{j+1}) , sendo que $v_i(t) = 0$ significa indicação semafórica vermelha e $v_i(t) = 1$, indicação semafórica verde; $u_i(t) \in \{0, 1\}$ é o valor do controle a ser aplicado no período subsequente, sendo 0 indicação verde e 1 indicação vermelha, também para esta variável e I é o número de interseções do sistema. O sistema assim descrito considera apenas dois estágios, podendo ser generalizado.

O tempo de vigência de um mesmo estágio, medido em termos de número de períodos, t_s sem mudança, é dado por:

$$w_i(t+1) = \begin{cases} w_i(t) + 1 & \text{se } u_i(t) = v_i(t) \\ 0 & \text{se } u_i(t) \neq v_i(t) \end{cases} \quad (4.2)$$

onde $w_i(t)$ é o número de períodos decorridos no estágio vigente.

Outro grupo de equações de estado descreve as variáveis que caracterizam as vias do sistema. O modelo vertical da fila, em uma faixa, tem evolução temporal dada por:

$$x_l(t+1) = \max\{0, x_l(t) + a_{l,1}(t) - s_l m_l(t)\} \quad (4.3)$$

onde $x_l(t)$ é a fila formada na linha de parada de uma faixa l ; $a_{l,1}(t)$ é número de veículos em velocidade livre que trafegam na primeira seção de uma faixa; s_l é a taxa média de descarga da fila, considerada conhecida e constante para todas as faixas e $m_l(t) \in \{0, 1\}$ é o controle da faixa em questão, sendo $m_l(t) = u_i(t)$ para as faixas arteriais e $m_l(t) = \overline{u_i(t)}$ para as faixas secundárias.

A discretização do modelo ocorre quando são criadas seções em uma faixa e o tempo necessário para atravessar uma seção é igual a um período de amostragem t_s , multiplicada pela velocidade (considerada uma dada velocidade livre de percurso constante). Estas seções são definidas por:

$$\mathcal{N}(l) = \frac{L_l}{v_l t_s} \quad (4.4)$$

onde $\mathcal{N}(l)$ é o número de seções de uma faixa; L_l é o comprimento de cada faixa, v_l é a velocidade livre de percurso e t_s é o tamanho do período de amostragem. $\mathcal{N}(l)$ é definido de tal forma que a passagem de uma seção para a seguinte é feita em exatamente um período amostral, como pode ser visto na Figura 4.1. A evolução da ocupação de veículos que trafegam em velocidade livre ao longo das seções de uma faixa, é dada por:

$$a_{l,j}(t+1) = a_{l,j+1}(t), j = 1, \dots, N_l - 1 \quad (4.5)$$

$$a_{l,N_l}(t+1) = a_{l,N_l+1}(t) + (1 - r_l)z_l(t) \quad (4.6)$$

$$a_{l,N_l+1}(t+1) = r_l z_l(t) \quad (4.7)$$

onde $z_l(t)$ é o fluxo de veículos no início da faixa, a montante da interseção; e $\mathcal{N}(l)$ é o número de períodos que um veículo gasta para atravessar uma faixa, sendo N_l a parte inteira e r_l a parte decimal deste período.

Os fluxos de veículos $z_l(t)$ são modelados de duas formas. Para faixas alimentadas por filas a montante, tem-se:

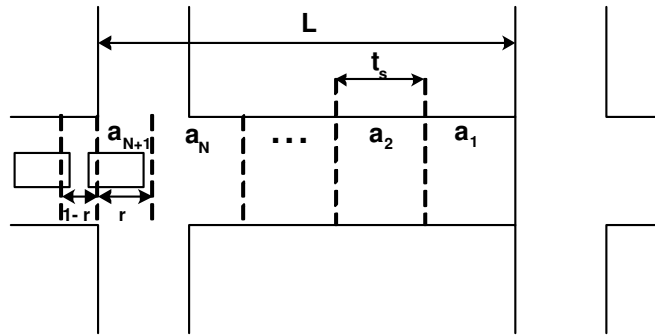


Figura 4.1: Seções de uma faixa.

$$z_l(t) = \sum_{l' \in U_l} p_{l'l} \times \min(x_{l'}(t) + a_{l',1}(t), s_{l'} m_{l'}(t)) \quad (4.8)$$

onde $l' \in U_l$, é a faixa pertencente a U_l que é o conjunto de faixas que alimentam um determinada faixa l ; $p_{l'l}$ é a proporção de veículos que convergem das filas para a faixa a jusante, considerada conhecida e constante. Para as faixas de entrada, $z_l(t)$ é simplesmente o resultado da contagem veicular.

O desempenho da rede é avaliado através do atraso acumulado. O modelo de tráfego ao fornecer a dinâmica das filas, permite avaliar o atraso acumulado no conjunto de todas as faixas l e em um horizonte $K + 1$:

$$d_l(t) = \frac{t_s}{2} \sum_{j=t+1}^{t+K+1} (x_l(j) + x_l(j-1)) \quad (4.9)$$

Pelo HCM 2000 [9], o atraso por veículo é dado por:

$$d_{vq} = l_s \frac{\sum_{i=1}^N V_{iq}}{V_{total}} 0.9 \quad (4.10)$$

onde l_s é o intervalo de contagem (s) (sugere-se entre 10 a 15 segundos); V_{iq} é o número de veículos na fila no intervalo de contagem i ; N é o número de intervalos de contagem ao longo de todos os ciclos observados e V_{total} é o número total de veículos chegando na aproximação durante o tempo total de observação (um determinado número de ciclo).

A atraso calculado pela equação 4.9 difere da equação 4.10 que considera o somatório das filas no tempo. Na equação 4.9 o cálculo considera o valor médio das filas e tem os valores da primeira fila e da última, no intervalo considerado, divididas por 2. Esta diferença não é significativa, se o intervalo de tempo total for maior do que o período. Também o cálculo é realizado da mesma forma para todas as aproximações, implicando em comparações válidas.

4.2 Algoritmo de Controle em Tempo Real

A idéia central do algoritmo de controle consiste em testar todas as possibilidades de controle $u_i(t)$ em um horizonte de tempo e escolher dentre todas aquela que produz o melhor desempenho. Tal teste é efetuado com auxílio do modelo de tráfego apresentado na seção 4.1.1, o qual, ao fornecer a dinâmica das filas, permite avaliar os índices de desempenho, que neste trabalho é medido pelo atraso acumulado no conjunto de todas as faixas l e em um horizonte $K + 1$, conforme equação (4.9).

O algoritmo de controle utiliza a técnica de horizonte deslizante [48] para tentar prever o comportamento futuro do tráfego e descobrir o melhor controle para o futuro. A técnica consiste em simular, entre dois instantes de controle, a evolução do sistema durante vários períodos. Ao final, tem-se o valor ótimo do controle para ser aplicado, o qual irá vigorar apenas durante o período subsequente. O processo é repetido para todos os períodos. As possibilidades criadas pelo horizonte deslizante podem ser representadas por uma árvore de decisão.

A árvore de decisão assim obtida é adequada para organizar a tarefa de testar todas as possibilidades. Ela é montada a partir das diferentes opções que uma decisão entre a indicação verde ou vermelha pode oferecer. Cada nova situação de tráfego criada tem um custo quantificado em atraso veicular. Um algoritmo de busca faz a tarefa de encontrar o caminho na árvore que resulte no menor custo. O resultado levará ao controle a ser aplicado no período subsequente.

Uma simplificação assumida no modelo de tráfego e no algoritmo de controle é o modelo de fila vertical. Este modelo é utilizado em vários outros modelos já citados [21], [26], [51].

Os aspectos e técnicas utilizadas serão discutidas mais detalhadamente nas seções seguintes.

4.2.1 Horizonte Deslizante

O procedimento de horizonte deslizante pode ser descrito da seguinte forma: a técnica exige o conhecimento das chegadas de veículos no período inicial. Para os períodos futuros do horizonte são utilizadas previsões das chegadas. Estas previsões são obtidas para todas as possibilidades de controle durante um horizonte de tempo K , com $k = 1, \dots, K$. Este procedimento utiliza um modelo de previsão, para este estudo, similar ao modelo de tráfego empregado. O tamanho do horizonte aplicado no algoritmo de controle em tempo real é $K = 8$, ou seja, $K = 8 t_s = 32s$.

O procedimento de otimização define o caminho de menor custo. Com este resultado define-se o controle aplicado em $k = 1$, neste caminho de menor custo. O controle em $k = 1$ é implementado no período subsequente e permanecerá vigente até que o procedimento seja repetido e novamente atualizado.

O resultado da aplicação do horizonte deslizante é representado por uma árvore de decisão. Esta árvore, para um semáforo de duas fases, é binária, pois, a decisão é tomada sobre uma variável discreta que pode assumir dois estados: a indicação verde ou vermelha. Os nós da árvore representam os estados do sistema. Suas arestas representam o custo da transição de um estado para outro. O custo da transição entre os estados é o atraso veicular, dado pela equação (4.9). Um exemplo de aplicação para um semáforo de dois estágios pode ser analisado na Figura 4.2.

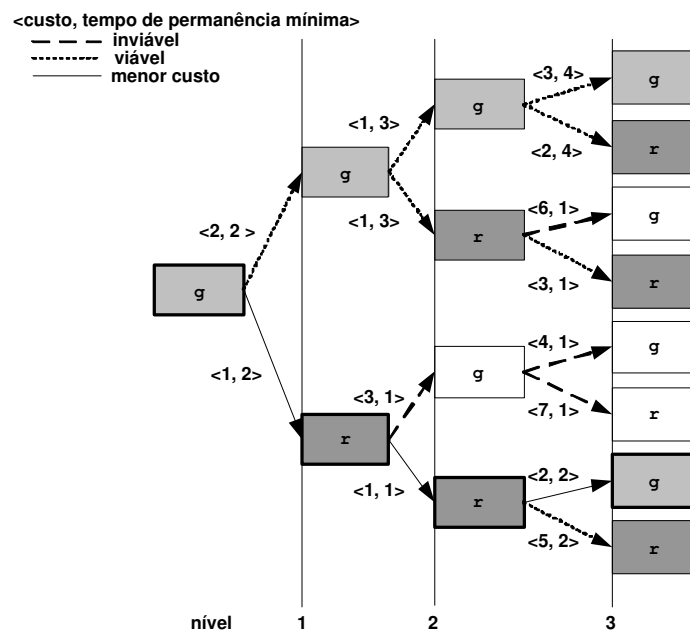


Figura 4.2: Árvore de busca - semáforo de dois estágios.

A Figura 4.2 mostra uma árvore completa para $K = 3$. A busca é realizada em profundidade, portanto é escolhido o ramo superior e percorrido o caminho até a última folha da árvore. Na figura, o caminho totaliza um custo de 6. A restrição de verde mínimo é cumprida, pois a indicação verde (g) é mantida durante todo o caminho. O segundo caminho a ser percorrido é o que testa a mudança para indicação vermelha (r) na metade inferior da árvore. E o caminho é o extremo inferior da árvore. Ao

chegar à última folha da árvore os valores de custo são comparados e opta-se por continuar a busca na metade da árvore onde foi encontrado o menor custo. Respeitando-se a restrição de verde mínimo e descartando-se os caminhos que, mesmo em folhas de nível inferior ao máximo, sejam maiores do que o menor custo atualmente encontrado, a busca continua até ser encontrado o menor custo no último nível da árvore. Assim, o caminho de menor custo igual a 4 aparece realçado na figura.

4.2.2 Predição de Chegadas

A predição de chegadas é realizada com o auxílio de um modelo de predição baseado no modelo de tráfego apresentado na seção 4.1.1. São consideradas chegadas o número de veículos que chegam nas vias que recebem veículos de um sistema externo ao sistema considerado. Na Figura ?? recebem fluxo externo as vias $\{1, 4, 5e6\}$. Algumas simplificações foram realizadas para tal aplicação:

1. o cálculo das chegadas é realizado, *off-line*. Com estes dados é montada uma tabela com a quantidade de veículos que chegam em cada nível do horizonte de tempo considerado, pela quantidade de arcos no sistema. Para os arcos internos, que não recebem fluxo externo, é realizado um cálculo com a soma das proporções de conversão de veículos dos arcos que descarregam fluxo no arco em questão;
2. o controle que vigora no período inicial do horizonte, isto é, as indicações semaforicas de cada interseção, é considerado constante para todo o horizonte. Esta consideração não implica em erros significativos pois, a cada tempo de amostragem é feita uma atualização de chegadas e refeito o cálculo de predição, corrigindo assim possíveis desvios.

A predição de chegadas é realizada (como descrito anteriormente no primeiro item) sempre anterior ao cálculo do controle realizado durante o horizonte deslizante. Durante o procedimento de horizonte deslizante a predição de entradas é realizada da mesma forma, no entanto o tipo de chegadas pode ser definido pelo operador, assim a quantidade de veículos que chegam no primeiro período do horizonte é conhecida, resultante das contagens nos detectores. Para os períodos seguintes pode ser utilizada:

1. constante;
2. a média dos períodos anteriores (no tamanho igual ao número de períodos do horizonte);
3. chegada nula de veículos.

4.2.3 Busca em Profundidade

O algoritmo de busca utilizado é o de busca em profundidade [15]. Esta escolha pode ser justificada pelos seguintes aspectos:

- simplicidade de implementação do método;
- a rapidez na obtenção de uma solução factível;
- pequeno consumo de memória durante a busca que é realizada em um caminho de cada vez, ao contrário, por exemplo da busca em largura, que precisa armazenar o valor do custo parcial de todas as folhas a cada nível que avança na árvore.

Por outro lado existe a desvantagem da explosão combinatória de caminhos que cresce exponencialmente com o tamanho do horizonte. Esta desvantagem poderia inviabilizar a aplicação do algoritmo dependendo do tamanho do horizonte. Assim, algumas modificações foram implementadas ao algoritmo com o objetivo de restringir a busca em todos os caminhos de uma árvore.

A principal modificação é permitir que a decisão a ser tomada (permanecer no estágio atual ou mudar) possa ser explorada alternativamente na árvore. Isto é, a busca é realizada, inicialmente, em uma direção da árvore, até atingir o último nível de profundidade, depois a busca retorna ao nó inicial e parte para a outra direção de busca. O valor do desempenho obtido no último nível da árvore, na direção inicial, é considerado o melhor até que uma busca em outro ramo qualquer (ao atingir o último nível) resulte em um valor de desempenho menor; neste ponto, invertem-se os papéis. Outra modificação é o descarte da busca em um determinado ramo ao ser obtido um valor de desempenho maior do que considerado ótimo até aquele momento da busca. Tal procedimento pode seguir até que seja realizada a busca em todos os ramos viáveis ou até que se atinja o limite de tempo disponível para o cálculo, quando então é adotada a decisão vencedora.

As modificações efetuadas também garantem que pelo menos as duas opções iniciais e distintas de controle são simuladas antes que um *dead-line* [23] possa ocorrer.

A busca na árvore tem outra restrição, além do *dead-line*, o tempo mínimo de duração da indicação semafórica dada por:

$$u_i(t) = v_i(t) \quad \text{se} \quad w_i(t) < w_{min} \quad (4.11)$$

A necessidade deste limite mínimo é justificada pelo custo da troca das indicações, baseado no tempo mínimo para que os veículos que estejam parados em uma fila possam se deslocar através da interseção. Esta restrição inibe algumas mudanças na temporização, poupando esforços computacionais na busca de um caminho de menor custo.

Para redes com um número maior de estágios em uma interseção, o algoritmo de busca em profundidade pode causar uma explosão combinatória, já que de cada nó parte um número de ramos igual ao número de estágios. A viabilidade de implementação do algoritmo existe desde que seja seguida uma seqüência de fases pré definida, o que torna a decisão em cada nó entre permanecer ou mudar para o próximo estágio.

4.2.4 Critério de Desempenho

A previsão das filas, ou seja, sua quantificação, é calculada através da equação de estados (4.3). O atraso médio acumulado no horizonte é dado pela equação (4.9).

Verifica-se que na equação (4.9) o atraso só será calculado para a interseção em interesse, conseqüentemente somente para as vias que convergem nesta interseção e no horizonte de tempo considerado. Para compor o critério de desempenho, este atraso é acrescido de um custo terminal. Adota-se aqui atraso incorrido aos veículos que permanecem na fila final, dado pelo tamanho destas filas dividido pela taxa média de descarga de fila:

$$\tilde{d}_l(t) = \frac{(x_l(t + K + 1))^2}{s_l} + \frac{t_s}{2} \sum_{j=t+1}^{t+K+1} (x_l(j) + x_l(j - 1)) \quad (4.12)$$

A introdução do custo terminal é uma tentativa de representar o comportamento de descarga da fila se o horizonte fosse considerado infinito. Vale salientar que o desempenho $\tilde{d}_l(t)$ só é aplicado ao algoritmo de controle, onde o horizonte considerado é pequeno. No modelo de tráfego o horizonte de simulação é maior, com ordem de grandeza de minutos ou horas, neste caso, o custo terminal pode ser desprezado, pois torna-se insignificante frente ao atraso acumulado no horizonte.

Algoritmo de Controle Descentralizado

gera-árvore ($v_e, \tilde{d}, d, p, p_{ot}$)

condição inicial: gera-árvore ($v_0, 0, +\infty, [v_0], 0$)

Se (passo (v_e) $\neq (K + 1)$)

Faça ($v_p \in \{F(v_e)\}$)

$q(v_e, v_p) = \text{calcula-atraso}$

$D = d + q(v_e, v_p)$

gera-árvore($v_p, D, \tilde{d}, p^+ \{v_p\}, p_{ot}$)

Senão

Se ($d + \sigma(v_e) < \tilde{d}$)

$\tilde{d} = d + \sigma(v_e)$

$p_{ot} = p$

onde:

- v_e é o nó inicial;
- v_p é nó final do passo seguinte;
- d é o custo de um caminho na árvore;
- σ é o peso de um nó;
- \tilde{d} é o custo mínimo otimizado;

- p é um caminho da árvore;
- p_{ot} é o caminho ótimo de menor custo;
- K é a extensão do horizonte;
- q é o custo calculado entre um nó inicial e um nó final;
- $\{F(v_e)\}$ é o conjunto de nós gerados de v_e , estes nós obedecem a restrição $w_i < w_{min}$ na mudança de indicação de fase;
- $calcula - atraso$ é uma função que calcula o custo conforme equação (4.12) e retorna o custo entre os nós da árvore;
- D é o custo acumulado a cada nível de profundidade avançado na árvore.

4.2.5 Coordenação Implícita

A arquitetura descentralizada do modelo e do algoritmo de controle dificulta uma estratégia de coordenação entre as interseções. No entanto, como no PRODYN e ALLONS-D, a coordenação ocorre de forma implícita devido a incorporação de dados das interseções a montante da interseção analisada. Ou seja, quando os cálculos de otimização são realizados para uma interseção, são conhecidos os dados de saída da interseção acima dela considerando constante para todo o horizonte a indicação semafórica vigente. Desta forma, a coordenação entre interseções é obtida implicitamente. Uma coordenação explícita também pode ser obtida através de restrições introduzidas ao algoritmo. Uma implementação deste tipo de coordenação, para o algoritmo de controle aqui proposto pode ser visto em [13].

4.3 Resultados de Simulações

Para avaliar a consistência do algoritmo proposto inicialmente será apresentada uma comparação com a estratégia TRANSYT, versão TRANSYT/10 [16]. O TRANSYT foi escolhido pois é considerado uma estratégia padrão para cálculo de tempos fixos. Os resultados de desempenho do algoritmo de controle descentralizado serão apresentados de forma a avaliar sua sensibilidade frente a diferentes distribuições de demanda e chegadas apresentadas no apêndice B. Baseado nos resultados também será apresentada a discussão de coordenação dos semáforos de modo a atender aos pelotões e descarga de filas.

4.3.1 Características da Rede de testes

Os testes foram realizados em uma rede de três interseções em um via arterial de sentido único de fluxo mostrada na Figura 4.3. O comprimento total da via arterial é de aproximadamente 760 m divididos em 3 faixas de aproximadamente 253 metros. as faixas secundários são de aproximadamente

350 metros. A velocidade dos veículos é considerada constante a um valor de 16,7 metros/segundo (60Km/h).

Serão apresentados resultados para diferentes instâncias de carregamento da rede apresentadas na Tabela 4.1. O carregamento considerado está abaixo do máximo grau de saturação, pois a concepção do algoritmo foi para controle de sistemas não-saturados. Diferentes distribuições de chegadas de veículos nas vias também foram consideradas. São quatro as distribuições: escada, exponencial, pulsada e constante. A forma como estes dados foram obtidos é apresentada no Apêndice B. Foram realizadas três simulações distintas para cada distribuição na geração de dados de chegadas. Foi feita uma média aritmética dos valores obtidos, para cada distribuição.

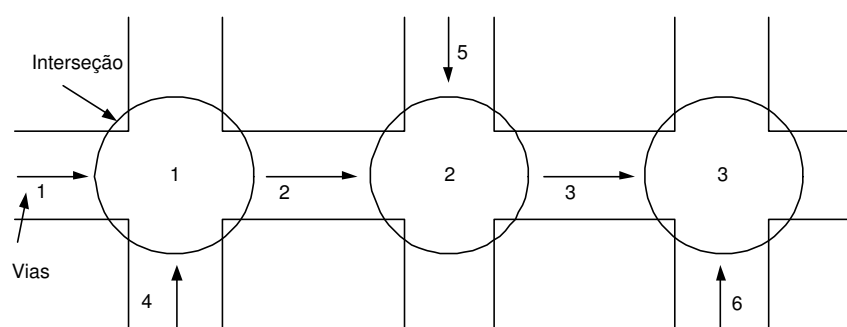


Figura 4.3: Rede de Tráfego para 3 interseções.

As amostras de tempo do algoritmo são de 4 segundos. As proporções de propagação são consideradas conhecidas e iguais a 90% na arterial e 10% das secundárias para a arterial. O fluxo de saturação é igual a 2 veículos por período de amostragem, também considerado conhecido e constante para todas as vias. A capacidade das vias é de 1800 veículos por hora (vph) (não está sendo considerado o tempo de amarelo e o tempo perdido). E o horizonte T de simulação $T = 30$ min.

Tabela 4.1: Proporção de fluxos nas faixas.

Caso	Arterial	Secundárias
desbalanceado	80%	20%
balanceado	55%	45%

Serão testadas duas situações de carregamento, considerando a soma das duas aproximações pois, a capacidade é considerada por interseção:

- alto - 90% da capacidade da interseção, ou seja, um total de 1620 veículos por hora (vph);
- médio - 60% da capacidade da interseção, ou seja, um total de 1080 vph.

4.3.2 Comparações com o TRANSYT

Na Tabela 4.2 são apresentados na primeira coluna o valor do atraso obtido diretamente da simulação realizada com o TRANSYT/10 em pcu-h/h; na segunda coluna são apresentados os valores

do modelo de tráfego descentralizado com tempos fixos (calculados pelo TRANSYT/10 e obtidos na mesma simulação do resultado da primeira coluna desta tabela), em veíc.-s/s, e na terceira coluna os valores obtidos diretamente do algoritmo de controle aplicado sobre o modelo de tráfego, em veíc.-s/s. Estas comparações foram realizadas com a distribuição de chegadas constante para satisfazer as especificações de entrada do TRANSYT/10. O mínimo tamanho de ciclo aceito pelo *software* TRANSYT/10 é de 30 segundos, que é um valor maior do que a média obtida com o algoritmo descentralizado.

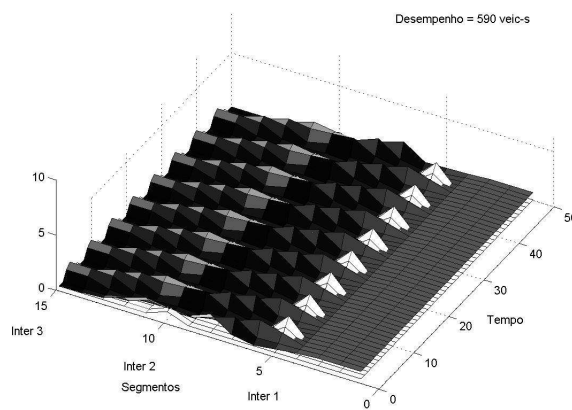
Tabela 4.2: Comparação entre valores de atraso com TRANSYT e estratégia descentralizada (pcu-h/h e pcu-s/s).

	TRANSYT	Fixo	Desc.
BA	27,0	19,2	5,2
DA	20,0	16,8	6,6
BM	9,6	10,7	2,9
DM	6,9	6,8	3,1

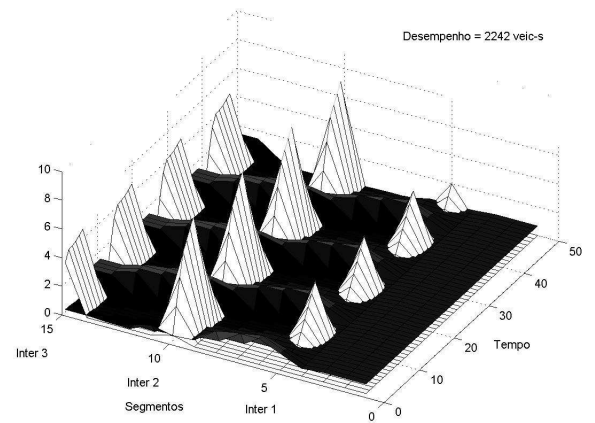
onde: BA é balanceado alto, DA é desbalanceado alto, BM é balanceado médio e DM é desbalanceado médio. A avaliação dos resultados apresentados na Tabela 4.2 permite verificar que os valores de atraso obtidos diretamente do TRANSYT são próximos aos valores obtidos através do algoritmo em tempo real (com os tempos fixos do TRANSYT). Era um resultado esperado pois apenas o modelo de tráfego e a forma de cálculo do desempenho são diferentes. Já o desempenho da rede obtido com o algoritmo em tempo real apresenta um ganho maior, o que também era esperado pois a otimização realizada procura responder à demanda verificada a cada amostra de tempo.

Outra comparação entre o desempenho obtido pelo TRANSYT e pelo algoritmo descentralizado pode ser verificado na Figura 4.4. É simulado um carregamento balanceado médio (BM) e desbalanceado médio (DM) e distribuição constante. Pode ser verificada a evolução dos pelotões, ao longo da faixa arterial, e a formação das filas nas interseções.

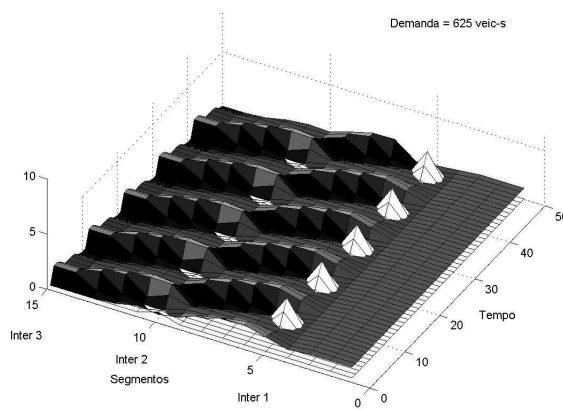
Na Figura 4.4 (a) é apresentado o comportamento do tráfego frente ao controle realizado pelo algoritmo em tempo real e em (b) o comportamento frente ao controle fixo para um carregamento médio balanceado. Pode-se verificar que em (b) os tempos de indicação verde e vermelha têm maior duração e que os ciclos são fixos. Foram testados valores de ciclo maiores e menores, no entanto o valor apresentado resulta em um melhor desempenho. Já em (a) os tempos de verde e vermelho são menores e adaptam-se a demanda verificada, implicando em uma menor formação de filas e, conseqüentemente em um melhor desempenho. As filas são verificadas como picos que ocorrem nas interseções e os pelotões como uma progressão nas seções da faixa, verificada em (a) e (b). Em (c) e (d) apresenta-se resultados para uma demanda média e desbalanceada. Verifica-se que para um fluxo maior na arterial o TRANSYT (d) proporciona uma melhoria de desempenho em relação a um fluxo balanceado entre as faixas (b). No entanto, ainda maior do que com os tempos calculados pelo algoritmo em tempo real (c).



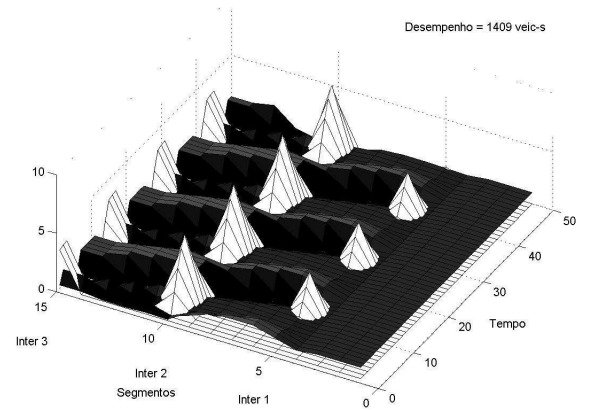
(a)



(b)



(c)



(d)

Figura 4.4: Comparação TRANSYT x algoritmo de descentralizado.

4.3.3 Desempenho do Algoritmo de Controle em Tempo Real

Nesta seção será apresentado o desempenho do algoritmo de controle em tempo real (como o algoritmo é baseado no modelo de tráfego de arquitetura descentralizada, a partir será identificado apenas como algoritmo descentralizado) para as quatro distribuições de chegadas e para as quatro situações de carregamento. O tempo de simulação é de 30 minutos e o resultado apresentado é dado em segundos. A distribuição em escada representa uma chegada de veículos randômica, onde são atribuídos valores inteiros entre 0 e 2, que representam os valores de chegada de veículos mínima e máxima, por período de amostragem. Os intervalos de distribuição são definidos para obtenção do valor total de chegadas de veículos por hora. A distribuição exponencial é uma distribuição contínua de probabilidades em que a ocorrência de veículos aumenta ou diminui constantemente. Neste trabalho foi utilizada para representar a chegada de veículos com um aumento constante no tempo. A distribuição constante é representada por uma chegada constante de veículos em quantidade suficiente para garantir a demanda definida. E a distribuição pulsada representa a alternância entre um valor de chegadas n , onde $0 < n \leq 2$ e *zero*. O valor 2 representa a capacidade máxima da via por período de amostragem. Os algoritmos de obtenção de valores para estas distribuições podem ser vistos no Apêndice B.

Tabela 4.3: Valores de atraso (s) obtidos com algoritmo descentralizado.

Caso	Escada	Exponencial	Pulsada	Constante
BA	12391	12002	8271	9426
DA	14352	13509	7812	11966
BM	5573	5498	3820	5193
DM	5050	5090	2026	5504

Os resultados apresentados na Tabela 4.3 mostram que para um alto carregamento, o atraso aumenta em relação a um médio carregamento, para todas as distribuições de chegadas. Um resultado importante que pode ser verificado aparece para a distribuição pulsada onde ocorre uma melhoria significativa no desempenho do sistema, resultado esperado, já que ocorrem períodos em que não chegam veículos e o algoritmo de controle adapta-se a esta situação. Isto sugere como perspectiva de estudos futuros um controle sobre os usuários que os mantivesse em pelotões.

São apresentados também resultados do algoritmo de controle em tempo real para duas situações distintas:

- Para uma única distribuição de demanda a comparação é realizada entre as diferentes distribuições de chegadas - esta avaliação tem como objetivo mostrar como o algoritmo reage a diferentes distribuições de chegadas mostrando assim sua robustez ao minimizar o atraso em diferentes situações de chegada.
- Para uma distribuição de chegadas aleatória a comparação é realizada entre diferentes demandas de fluxo - o objetivo deste teste é também avaliar a sensibilidade do algoritmo, em minimizar o atraso e proporcionar a progressão de pelotões agora frente a diferentes demandas de fluxo.

A Figura 4.5 apresenta os resultados para diferentes distribuições de chegadas e um fluxo desbalanceado médio. Esta comparação proporciona a análise de como o algoritmo de controle reage às diferentes situações de tráfego.

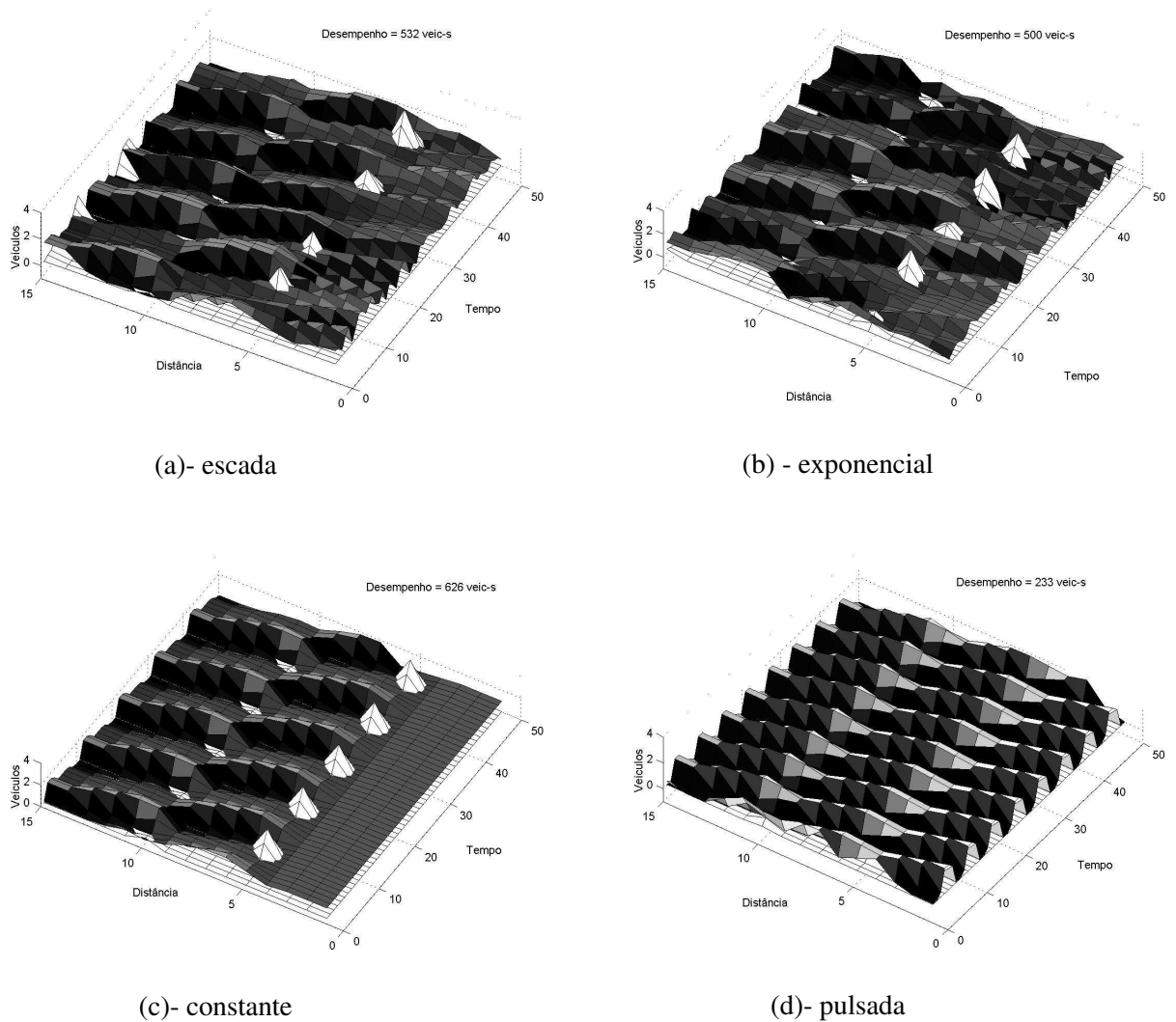
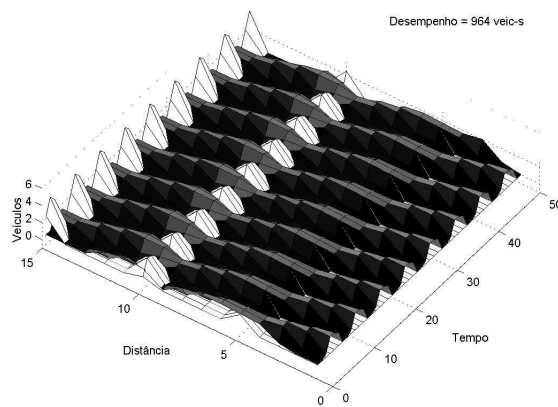


Figura 4.5: Comparação entre diferentes distribuições de chegadas.

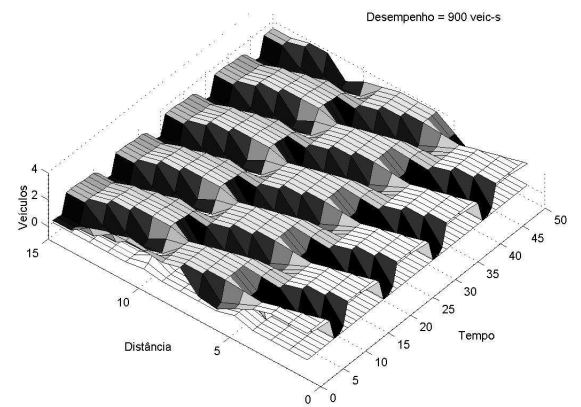
Pode ser verificado na Figura 4.5 que a maior formação de filas ocorre para as distribuições escada e exponencial (a) e (b), já que a predição do comportamento destas distribuições é dificultada pela natureza das chegadas. Já com chegadas constante e pulsada o controle consegue rapidamente adequar-se e proporcionar um melhor desempenho.

A Figura 4.6 apresenta a formação de filas na via arterial para a segunda situação. Pode ser verificado a distribuição de veículos nas seções da via e as filas nas interseções, para os quatro casos de carregamento.

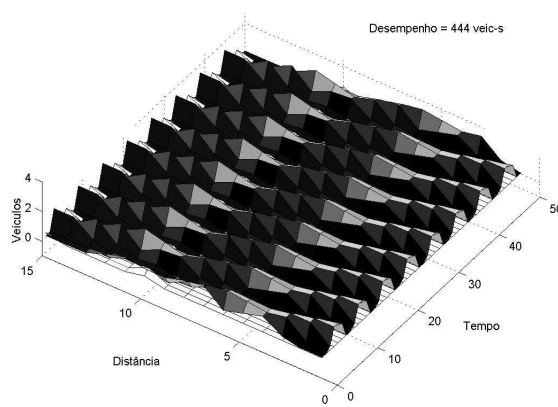
Pode-se verificar na Figura 4.6 (a) e (c) que, para uma distribuição de fluxo balanceado, ocorre uma queda de desempenho, (aumento das filas) em relação a distribuição de fluxo desbalanceada (b) e (d).



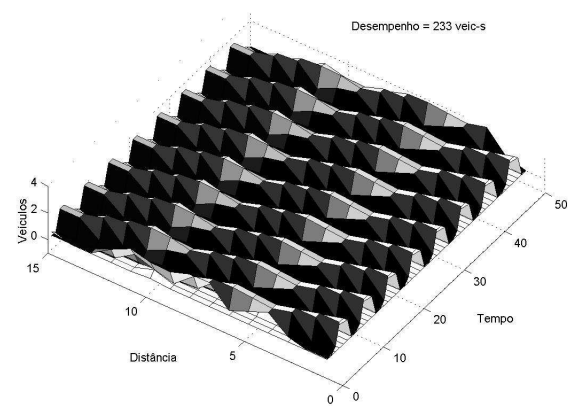
(a) balanceado alto



(b) desbalanceado alto



(c) balanceado médio



(d) desbalanceado médio

Figura 4.6: Comparação entre diferentes distribuições de fluxo para chegadas pulsadas.

4.4 Conclusões

Neste Capítulo foi apresentado o modelo de tráfego que serviu de suporte para simulação de uma rede de tráfego e base para o desenvolvimento do algoritmo de controle. Foram realizadas simulações para avaliação do desempenho do algoritmo proposto frente ao TRANSYT e mostrado o comportamento do tráfego frente ao controle obtido com o algoritmo descentralizado em tempo real.

Os resultados obtidos permitem afirmar que o algoritmo de controle descentralizado tem melhor desempenho do que o TRANSYT.

Uma limitação do algoritmo descentralizado é que os testes realizados ocorreram para uma rede de tráfego com apenas duas fases. Desta forma, sua representação através da árvore de busca é binária. A dificuldade de aplicação deste mesmo algoritmo para redes com um número maior de fases é a explosão combinatória, pois o número de ramos que saem de cada nó deve ser igual ao número de fases.

É importante também registrar que o modelo de tráfego utilizado para estimação de atrasos ao longo do horizonte é o mesmo que simula a rede de tráfego real (cidade). Portanto, como perspectiva, é preciso testar o desempenho do algoritmo de controle descentralizado sobre micro-simulação.

Ainda na predição de chegadas foram consideradas chegadas pontuais, ou seja, a cada intervalo de tempo ocorre uma chegada qualquer de veículos ou não. Um aprimoramento que faria o modelo aproximar-se da realidade seria a consideração de chegadas baseadas em probabilidades. Para cada intervalo de tempo seria definido um vetor cujos elementos seriam probabilidades dos valores correspondentes aos índices deste vetor. Assim, o elemento zero corresponderia a probabilidade de chegadas nulas, o elemento um a probabilidade de chegada de um veículo por intervalo de tempo e assim continuamente. Um exemplo de predição de chegadas baseado em probabilidades foi aplicado em [21].

Capítulo 5

Modelo de Tráfego em Programação Matemática

5.1 Introdução

No capítulo 4 foi apresentado o modelo de tráfego desenvolvido para este trabalho. Este modelo será representado através de uma formulação matemática para otimização. Esta formulação tem duas finalidades, a primeira, para servir de padrão para estratégias de controle descentralizadas e, a segunda, para examinar a viabilidade dos pacotes de otimização para uso em tempo real. O modelo leva em consideração a influência das ações de controle mútuas entre interseções vizinhas, a dinâmica do fluxo do tráfego e restrições ao controle. A natureza linear inteira mista do modelo [60] permite usar a teoria de programação inteira [62] com a vantagem de utilização de algoritmos e *software* numéricos tais como *Xpress-MP* [33] e *ILOG Cplex* [4] que podem encontrar soluções globais ótimas para o problema de controle semafórico. Estes *software* aceitam algumas linguagens de modelagem tais como AMPL [24] e Mosel [33]. Tais linguagens procuram separar o modelo dos dados, permitindo assim que um modelo possa ser utilizados na resolução de diferentes instâncias de uma mesma classe de problemas. Elas também são responsáveis pela parte de pré-processamento e interface com os algoritmos de otimização.

O modelo de tráfego descrito em programação matemática com o objetivo de buscar um valor de desempenho ótimo global pode ter um custo computacional proibitivo. Ainda assim, torna-se interessante comparar os resultados globais com aqueles obtidos por soluções heurísticas e algoritmos aproximativos, como usados no Capítulo 4. A qualidade da solução pode ser avaliada por comparação à solução obtida com o modelo global. Esta comparação pode ser feita por violação de restrições e/ou comparação de valores da função objetivo.

5.2 Modelo para Programação Matemática

A dinâmica do sistema de tráfego, baseada no modelo descrito na seção 4.1.1, terá sua topologia modelada por um grafo $G = (V, E)$ cujos nós representam as interseções das vias e os faixas representam as vias. Os nós subdividem-se em dois tipos: os que representam as interseções controladas e os limítrofes, que representam interseções sobre as quais não há autoridade de controle. Estes nós limítrofes são estabelecidos como fronteiras da rede analisada com o restante do sistema de tráfego. Na Figura 5.1, os nós-interseção, círculos escuros, são representados por $I = \{1, 2, 3\}$, e os nós-fronteira, círculos claros, são representados por $B = \{4, \dots, 7\}$, onde $V = I \cup B$. Um exemplo de aplicação de modelagem AMPL, em uma rede de dimensão maior, com duplo sentido de fluxo está apresentado em detalhes, no apêndice A.

A rede em consideração é denotada por Γ . O problema P para operação de Γ pode ser aproximado pelo modelo de programação matemática apresentado a seguir. O objetivo é minimizar o atraso total acumulado, de acordo com a seguinte função objetivo:

$$P : \quad \text{Minimizar} \quad f = \frac{t_s}{2} \sum_{i \in I} \sum_{l \in L_i} \sum_{t=1}^{T-1} \{x_l(t) + x_l(t+1)\} \quad (5.1)$$

onde $L_i \subseteq E$ é o subconjunto das faixas que chegam na interseção i , $\cup_{i \in I} L_i = E$, $L_m \cap L_j = \emptyset, \forall m \neq j$ e t_s é o tempo de amostragem sendo $t_s = t_{j+1} - t_j$, para $j = 1, \dots, T-1$. Além disso, T é o número de períodos do horizonte de simulação considerado e $x_l(t)$ é o número de veículos estacionários na linha de parada da faixa l . O estado inicial da rede de tráfego é dado por:

para cada $l \in L_i$:

$$x_l(0) = x_l^0 \quad (5.2)$$

$$a_{l,j}(0) = a_{l,j}^0, \quad j = 1, \dots, N(l) + 1 \quad (5.3)$$

onde $x_l^0(t)$ é o número inicial de veículos parados e $a_{l,j}^0$ é o número inicial de veículos em velocidade livre na j ésima seção da faixa l , conforme equação (4.5).

O número de seções de uma faixa l , $\mathcal{N}(l)$ é igual ao número de intervalos de controle t requeridos para atravessar o comprimento da faixa l , sendo dado pela equação (4.4). Note que $\mathcal{N}(l)$ é, em geral, não inteiro. Então, existirá um número inteiro de seções que são atravessadas em um intervalo de controle e uma seção de comprimento fracionário.

O carregamento das filas e suas descargas são calculadas para cada $i \in I$ e são formulados como segue:

para cada $l \in L_i, t = 0, \dots, T-1$:

$$x_l(t+1) \geq x_l(t) + a_{l,1}(t) - y_l(t) \quad (5.4)$$

$$y_l(t) \leq x_l(t) + a_{l,1}(t) \quad (5.5)$$

$$y_l(t) \leq s_l m_l(t) \quad (5.6)$$

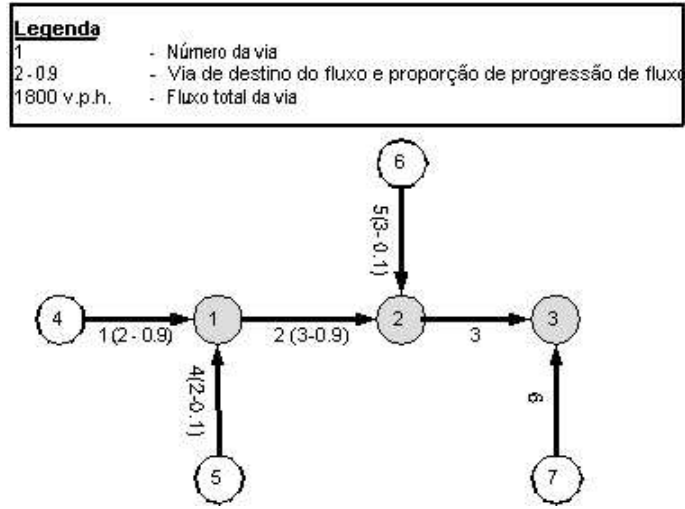


Figura 5.1: Grafo da rede de tráfego com 3 interseções.

onde $a_{l,1}(t)$ é o número de veículos em movimento através da primeira seção do $l^{\text{ésima}}$ faixa (seção imediatamente anterior à linha de parada); $m_l(t)$ assume o valor 1 se o tráfego da faixa l pode fluir durante o $t^{\text{ésimo}}$ período, e 0 caso contrário; s_l é a taxa de descarga máxima da faixa l ; $y_l(t)$ é o número de veículos que partem da faixa l , durante o $t^{\text{ésimo}}$ período. Note que $y_l(t) = 0$ se $m_l(t) = 0$ e $y_l(t) = \min\{s_l, x_l(t) + a_{l,1}(t)\}$ se $m_l(t) = 1$.

A progressão ao longo das faixas que levam à interseção i é dada pelas seguintes restrições:

$$\begin{aligned} &\text{para cada } l \in L_i, \quad t = 0, \dots, T-1, \quad j = 1, \dots, N(l)-1 : \\ &a_{l,j}(t+1) = a_{l,j+1}(t) \end{aligned} \quad (5.7)$$

O volume de tráfego a montante da faixa l é dado pelas equações (4.5) e (4.6, sendo reescritas na forma de programação matemática como:

$$\begin{aligned} &\text{para cada } l \in L_i \cap W_y, \quad t = 0, \dots, T-1 : \\ &a_{l,N(l)}(t+1) = a_{l,N(l)+1}(t) + (1-r_l)z_l(t) \end{aligned} \quad (5.8)$$

$$a_{l,N(l)+1}(t+1) = r_l z_l(t) \quad (5.9)$$

$$z_l(t) = \sum_{l' \in U_l} p_{l',l} y_{l'}(t) \quad (5.10)$$

onde $W_y \subseteq E$ é o subconjunto de faixas que recebem fluxo de outras faixas; $N(l)+1$ é o número de seções da faixa l ; r_l é a parte fracionária da faixa l que ocupa a $(N(l)+1)^{\text{ésima}}$ seção enquanto $(1-r_l)$ é a fração de veículos que entram na $N(l)^{\text{ésima}}$ seção em um intervalo; $U_l \subseteq E$ é o subconjunto de faixas cujos veículos podem entrar na faixa l , $l \in W_y$; $z_l(t)$ é o número de veículos provenientes de outras faixas que entram na faixa l durante o período t , faixas estas correspondendo a $l' \in U_l$; e $p_{l',l}$ é a fração de veículos que trafegam da faixa l' para a faixa l .

As equações (5.8-5.9) e (5.11-5.12) são necessárias em vista do resultado fracionário da equação (4.4).

Exemplificando para a rede da Figura 5.1, o volume que chega das faixas a montante aa faixa $l = 2$ é obtido fazendo-se $U_2 = \{1, 4\}$ com $W_y = \{2, 3\}$.

Para modelar o tráfego que entra na rede através de seus limites externos, se fazem necessárias as seguintes restrições:

para cada $l \in L_i \cap W_n, t = 1, \dots, T$:

$$a_{l,N(l)}(t) = a_{l,N(l)+1}(t-1) + (1 - r_l)a_l^t \quad (5.11)$$

$$a_{l,N(l)+1}(t) = r_l a_l^t \quad (5.12)$$

onde W_n é o subconjunto de faixas que recebem tráfego externo; note que $W_y \cap W_n = \emptyset$ e $W_y \cup W_n = E$ e a_l^t é uma constante que representa a quantidade prevista de veículos que chegam na faixa $l \in W_n$. Para a rede da Figura 5.1, $W_n = \{1, 4, 5, 6\}$.

O estado da indicação corrente, em termos do número de verdes consecutivos garantidos para a faixa l , iniciando no tempo $t - 1$, é dado por:

para cada $l \in L_i, t = 0, \dots, T - 1$:

$$q_l(t+1) \leq q_l(t) + m_l(t) \quad (5.13)$$

$$q_l(t+1) \geq q_l(t) + m_l(t) - M[1 - m_l(t)] \quad (5.14)$$

$$q_l(t+1) \leq Mm_l(t) \quad (5.15)$$

$$g_l m_l(t+1) \geq g_l - q_l(t) - g_l M[1 - m_l(t)] \quad (5.16)$$

onde $q_l(t)$ é o número de verdes consecutivos garantidos para a faixa l no período imediatamente precedente a t ; q_l^0 é o número inicial de verdes consecutivos concedidos aa faixa l e imediatamente anteriores ao período $t = 0$; M é um número inteiro positivo suficientemente grande, onde $M = T + 1$ satisfaz; e $g_l \in \mathbb{Z}_+$ é o tempo de verde mínimo expresso em amostras de tempo. As condições iniciais para indicação semafórica formam um grupo de restrições adicionais:

para cada $l \in L_i$:

$$q_l(0) = q_l^0 \quad (5.17)$$

A resolução do conflito é assegurada pelas seguintes restrições:

para cada $l \in L_i, t = 0, \dots, T - 1$:

$$m_l(t) \in \{0, 1\} \quad (5.18)$$

para cada $t = 0, \dots, T - 1$:

$$\sum_{l \in L_i} m_l(t) \leq 1 \quad (5.19)$$

as quais asseguram que $m_l(t)$ assume valores binários e, para um dado tempo t , somente um faixa

tem permissão de passagem.

Assume-se que todas as variáveis são números reais positivos, com exceção de $m_l(t)$ que assume valores binários. O problema P pode ser resolvido ao encontrar-se solução para (5.1) sujeito às restrições (5.2) - (5.19), sendo que as restrições são parametrizadas para cada interseção i e instanciadas para todo $i \in I$.

O problema P pode ser inserido na classe de problemas de programação linear inteira mista, sendo que as variáveis inteiras são as indicações semaforicas. Duas vantagens desta formulação são:

- a obtenção de limites inferiores através da solução da relaxação linear de P , dessa forma permitindo a avaliação da qualidade de soluções aproximadas ou heurísticas;
- o emprego de algoritmos de otimização para encontrar soluções ótimas de P e, possivelmente, soluções sub-ótimas quando um procedimento de enumeração implícita (p.ex., *branch-and-bound* [62]) sofrer uma interrupção prematura.

Em princípio, se a dinâmica do tráfego fosse determinística e se existissem recursos computacionais abundantes, a solução de P forneceria tempos de chaveamento ótimos para todos os semáforos no horizonte de tempo T . Com relação aos dois impedimentos:

- *dinâmica não determinística* - o modelo de controle preditivo [46] e [54], instancia uma série $\{P_t : t = 0, 1, \dots\}$ de problemas, onde cada problema P_t aproxima P sobre um horizonte deslizante [47] limitado e faz uso de realimentação periódica para compensar os erros nas predições, que tendem a tornar-se imprecisas com o expansão do horizonte;
- *recursos computacionais abundantes* - sozinho, o modelo de controle preditivo não proporciona uma solução prática; assim, faz-se necessária a decomposição de P_t em um conjunto de sub-problemas $\{P_{t,i} : i \in I\}$, um para cada interseção, o que converte a série $\{P_t\}$ em uma série de conjuntos de problemas $\{\{P_{t,i}\}\}$.

O algoritmo em tempo real apresentado no Capítulo 4 é comparável ao modelo de programação matemática quanto ao instanciamento de uma série de problemas $\{P_t : t = 0, 1, \dots\}$ quando é aplicada a técnica de horizonte deslizante para prever as chegadas e conseqüentemente o comportamento futuro do tráfego. Quanto à decomposição de P em sub-problemas, a similaridade se estabelece na arquitetura descentralizada do algoritmo de controle quando trata localmente a otimização do controle das interseções.

5.3 Resultados

Os resultados apresentados nesta seção são obtidos através de testes efetuados com a rede de tráfego da Figura 4.3, mantidas as mesmas características apresentadas no Capítulo 4. Somente o tempo

Tabela 5.1: Comparação do atraso total (s) entre algoritmo descentralizado (D) e modelo global (G). Simulação realizada para $25t_g$.

	Escada		Exponencial		Pulsada		Constante	
	G	D	G	D	G	D	G	D
BA	517	581	608	665	474	563	512	512
DA	639	746	841	935	406	418	623	674
BM	295	343	249	290	182	221	286	289
DM	216	230	234	267	107	107	253	304

de simulação é alterado para 25 amostras de tempo (4 segundos cada), por restrições de eficiência computacional do modelo de programação matemática.

Na Tabela 5.1 são apresentados os valores de atraso total obtidos com o algoritmo descentralizado e com o modelo global para programação matemática. Os resultados mostram, nas colunas G , os valores de ótimos globais obtidos com o modelo de programação matemática aplicado ao *software* ILOG Cplex AMPL 9.0. e na coluna D os valores obtidos com o algoritmo descentralizado de controle em tempo real, discutidos no Capítulo 4. Os resultados mostram um melhor desempenho do modelo global em relação ao modelo descentralizado. Isto se deve ao modelo global levar em consideração a influência das ações de controle mútuas entre interseções vizinhas e o modelo descentralizado sofrer simplificações heurísticas para permitir sua aplicação em tempo real. No entanto, pode-se verificar que em alguns casos como para distribuição de chegadas pulsada desbalanceada média e chegada constante balanceada alta os resultados de otimização com o algoritmo descentralizado chegaram ao resultado ótimo global.

Na Tabela 5.2 pode-se verificar a percentagem de melhoria do desempenho encontrado com o modelo global sobre o algoritmo descentralizado. Os valores variam de um mínimo de 1,1% a um máximo de 21,1%, sendo que em dois casos o algoritmo descentralizado atinge valores de ótimo global. Desta forma pode-se afirmar que em função das características de tráfego simuladas o ganho do modelo global sobre o descentralizado não é relevante, no entanto, para chegadas pulsadas com distribuição balanceada entre as faixas secundários e arteriais o algoritmo descentralizado distanciou-se do valor de ótimo global, o inverso acontecendo para uma situação de distribuição de tráfego desbalanceada. A proporção apresentada na Tabela 5.2 foi calculada como:

$$\text{Proporção} = \frac{D - G}{G} 100 \quad (5.20)$$

Tabela 5.2: Proporção de melhoria de desempenho do modelo global \times algoritmo descentralizado.

	Escada	Exponencial	Pulsada	Constante
BA	12,5	9,5	18,8	0,0
DA	16,8	11,2	3,0	8,3
BM	16,3	16,5	21,1	1,1
DM	6,3	14,0	0,00	20,1

Na Figura 5.2 pode ser verificada a progressão dos pelotões para um carregamento balanceado

médio (BM) com uma distribuição de chegadas em escada. Verifica-se que o comportamento dos pelotões e a formação de filas são próximos entre os resultados das duas estratégias. A progressão dos pelotões pode ser identificada pela formação de vales, caracterizando a descarga das filas formadas.

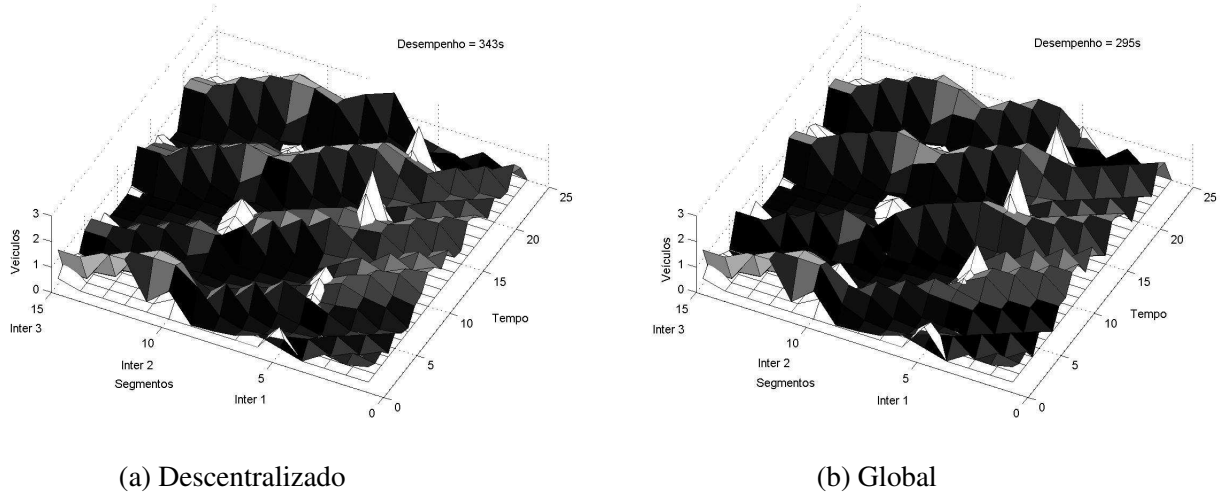


Figura 5.2: Comparação algoritmo descentralizado (a) × modelo global (b), carregamento desbalanceado médio, chegada em escada.

5.4 Extensões ao Modelo de Programação Matemática

O modelo matemático (global) apresentado é uma descrição em programação matemática do modelo de tráfego apresentado no Capítulo 4. As limitações do modelo de tráfego são repetidas no modelo global. A superação de algumas limitações são apresentadas como possíveis extensões ao modelo global e conseqüentemente ao modelo de tráfego e ao algoritmo de controle descentralizado em tempo real. Estas extensões não foram testadas e, portanto, os resultados dos testes realizados não incluem estas características.

5.4.1 Restrição de Verde Máximo

A restrição de verde máximo torna-se necessária quando faixas com uma demanda de veículos pequena em relação a faixa em conflito, podem não ter formação de fila suficiente para receber a indicação verde. Desta forma, uma pequena quantidade de veículos poderia ficar retida por um tempo abusivo caso não exista uma restrição de verde máximo.

Seja g_l^{max} o número máximo de indicações semafóricas verdes contínuas que podem ser atribuídas à faixa l . Pode-se implementar esta restrição da seguinte forma:

$$\begin{aligned} &\text{para cada } i \in I, l \in L_i, t = 0, \dots, T - 1 : \\ &q_l(t) \leq g_l^{max} \end{aligned} \tag{5.21}$$

Nestas circunstâncias, pode-se substituir a constante M em (5.14) e (5.16) por g_l^{max} .

5.4.2 Restrição de Vermelho Máximo

A restrição de vermelho máximo se aplica ao mesmo caso da restrição de verde máximo.

Seja r_l^{max} o número máximo de indicações vermelhas que o faixa l pode receber continuamente. Pode-se introduzir esta restrição ao modelo da seguinte forma:

para cada $i \in I, l \in L_i, t = 0, \dots, T - 1$:

$$q_l^r(t+1) \leq q_l^r(t) + (1 - m_l(t)) \quad (5.22)$$

$$q_l^r(t+1) \geq q_l^r(t) + (1 - m_l(t)) - r_l^{max} m_l(t) \quad (5.23)$$

$$q_l^r(t+1) \leq r_l^{max}(1 - m_l(t)) \quad (5.24)$$

onde $q_l^r(t)$ é o número de vermelhos consecutivos atribuídos aa faixa l no período imediatamente precedente a t .

5.4.3 Coordenação entre semáforos adjacentes

Supondo que a coordenação deve ser garantida entre a faixa l_k da interseção i_k ($l_k \in L_{i_k}$) e a faixa l_j da interseção i_j ($l_j \in L_{i_j}$), conforme Figura 5.3. Para uma indicação verde para a faixa l_k no instante t , $m_{l_k}(t) = 1$, deseja-se que a via l_j receba indicação de verde durante todo intervalo de tempo $[t + \Delta t_{k,j}^{min}, t + \Delta t_{k,j}^{max}]$ o que implica $m_{l_j}(t') = 1$ para todo $t' \in \{t + \Delta t_{k,j}^{min}, t + \Delta t_{k,j}^{min} + 1, \dots, t + \Delta t_{k,j}^{max}\}$

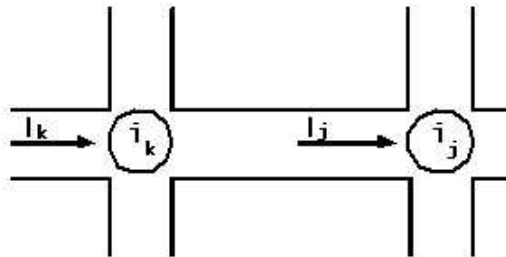


Figura 5.3: Rede para coordenação.

A especificação acima pode ser transcrita em programação matemática:

para todo $t = 0, \dots, T - 1$:

para $t' = t + \Delta t_{k,j}^{min}, \dots, \min\{t + \Delta t_{k,j}^{max}, T - 1\}$

$$m_{l_j}(t') \geq m_{l_k}(t) \quad (5.25)$$

Caso seja desejável que a indicação verde na via l_j ocorra em pelo menos um instante do intervalo

$[t + \Delta t_{k,j}^{min}, t + \Delta t_{k,j}^{max}]$, então basta substituir a equação (5.25) por:

$$\begin{aligned} &\text{Para todo } t = 0, \dots, T - 1 : \\ &\sum_{t' = t + \Delta t_{k,j}^{min}}^{\min\{t + \Delta t_{k,j}^{max}, T - 1\}} m_{l_j}(t') \geq m_{l_k}(t) \end{aligned} \quad (5.26)$$

5.4.4 Aproximação de Filas Horizontais

Para obtenção de um modelo de filas horizontais, a variável $a_{l,j}(t)$ passa a modelar ocupação da seção j da via l durante o período t . A variável $x_l(t)$ deixa de existir e introduz-se a variável $q_{l,j}(t)$, $j = 1, \dots, N(l)$, tal que $q_{l,j}(t)$ corresponde ao número de veículos que saem da seção $j + 1$ e atingem a seção j durante o período t .

1. **Restrição de capacidade de seção:** O modelo de fila vertical não permite a identificação do final espacial da fila formada. Uma aproximação de filas horizontais com a restrição de capacidade das seções de uma faixa torna possível a verificação deste limite. Esta restrição determina quando uma faixa tem sua capacidade atingida pela localização de veículos parados nas seções da faixa ou veículos em velocidade de cruzeiro antes de uma parada.

$$\begin{aligned} &\text{para todo } l \in E, t = 0, \dots, T : \\ &a_{l,1}(t) \leq a_{l,1}^{max,parado} \end{aligned} \quad (5.27)$$

$$\begin{aligned} &\text{para todo } l \in E, j = 2, \dots, N(l) + 1, t = 0, \dots, T : \\ &a_{l,j}(t) \leq a_{l,j}^{max,parado} \end{aligned} \quad (5.28)$$

$$a_{l,j}(t) \leq a_{l,j}^{max,cruzeiro} + \alpha_{l,j} \sum_{i=1}^{j-1} a_{l,i}(t) \quad (5.29)$$

onde:

- $a_{l,j}^{max,parado}$ é uma constante com o número máximo de veículos parados que podem ser acomodados na seção j da via l ;
- $a_{l,j}^{max,cruzeiro}$ é uma constante com o número máximo de veículos que podem ocupar a seção j da via l em velocidade de cruzeiro;
- $\alpha_{l,j}$ é um fator de ajuste (constante) que aumenta a capacidade de ocupação da seção j , conforme o acúmulo de veículos à jusante da seção j , ou seja, à medida que os veículos reduzem a velocidade, a ocupação pode ser acrescida. Este valor deve ser escolhido de tal forma que:

$$\alpha_{l,j} = \frac{a_{l,j}^{max,parado} - a_{l,j}^{max,cruzeiro}}{\sum_{i=1}^{j-1} a_{l,i}^{max,parado}} \quad (5.30)$$

$$\text{Se } a_{l,j}^{max,parado} = A_{max}, \forall j$$

$$\alpha_{l,j} = \frac{1}{A_{max}(j-1)} (A_{max} - a_{l,j}^{max,cruzeiro}) \quad (5.31)$$

A Figura 5.4 representa uma aproximação dos valores obtidos com as equações (5.30) e (5.31).

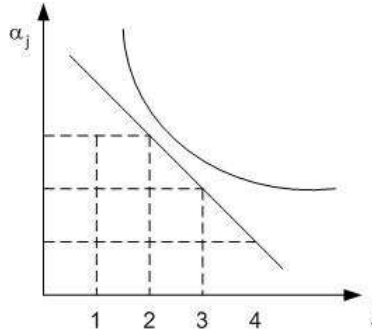


Figura 5.4: Valores hipotéticos de α_j .

2. **Restrição de fluxo** A restrição de fluxo permitirá a entrada de veículos nas seções de uma faixa somente quando houver capacidade nesta seção para entrada de novos veículos.

Para as seções inteiras da faixa l , considera-se:

$$\begin{aligned} &\text{para } l \in E, j = 2, \dots, N(l) - 1, t = 0, \dots, T - 1 : \\ &a_{l,j}(t+1) = a_{l,j}(t) + q_{l,j}(t) - q_{l,j-1}(t) \end{aligned} \quad (5.32)$$

Para a seção da faixa l imediatamente após a linha de parada, considera-se:

$$\begin{aligned} &\text{para } l \in E, t = 0, \dots, T - 1 : \\ &a_{l,1}(t+1) = a_{l,1}(t) + q_{l,1}(t) - y_l(t) \end{aligned} \quad (5.33)$$

A equação (5.4) é removida. A equação (5.5) é substituída por:

$$y_l(t) \leq a_{l,1} + q_{l,1}(t) \quad (5.34)$$

E a equação 5.6 deve ser mantida:

$$y_l(t) \leq s_l m_l(t)$$

Para as seções da faixa l não inteiras, tem-se:

$$\begin{aligned} &\text{para } l \in E, t = 0, \dots, T - 1 : \\ &a_{l,N(l)}(t+1) = a_{l,N(l)}(t) - q_{l,N(l)-1}(t) + (1 - r_l)z_l(t) + q_{l,N(l)}(t) \end{aligned} \quad (5.35)$$

$$a_{l,N(l)+1}(t+1) = a_{l,N(l)+1}(t) - q_{l,N(l)}(t) + r_l z_l(t) \quad (5.36)$$

3. Dinâmica de fluxo

para todo $l \in E, j = 2, \dots, N(l) - 1, t = 0, \dots, T - 1$:

$$q_{l,j}(t) = \text{Min}\{a_{l,j+1}(t), a_{l,j}^{max, cruzeiro} + \alpha_{l,j} \sum_{i=j-1}^1 a_{l,i}(t) - a_{l,j}(t) + q_{l,j-1}(t)\} \quad (5.37)$$

para todo $l \in E, t = 0, \dots, T - 1$:

$$q_{l,1}(t) = \text{Min}\{a_{l,2}(t), a_{l,1}^{max, parado} - a_{l,1}(t) + y_l(t)\} \quad (5.38)$$

$$q_{l,N(l)}(t) = \text{Min}\{a_{l,N(l)+1}(t), a_{l,N(l)}^{max, cruzeiro} + \alpha_{l,N(l)} \sum_{i=N(l)-1}^1 a_{l,i}(t) - a_{l,N(l)}(t) + q_{l,N(l)-1}(t)\} \quad (5.39)$$

A função $\text{Min}\{a, b\}$ pode ser expressa em programação matemática usando variáveis binárias e algumas igualdades e desigualdades lineares.¹

5.4.5 Discussão

As extensões das seções 5.4.1, 5.4.2 e 5.4.3, não introduzem dificuldades intrínsecas ao modelo, ou seja, não são introduzidas variáveis binárias ou discretas adicionais. Todas as restrições incluídas são lineares. Somente para a extensão de restrição de verde máximo, são introduzidas variáveis adicionais contínuas (g_l^{mx}).

Já para a extensão da seção 5.4.4 é adicionado um número significativo de variáveis binárias para expressar a função não linear $\text{Min}\{a, b\}$ em programação matemática.

A implementação das extensões descritas não foi realizada, são, no entanto, considerados para trabalhos futuros.

5.5 Conclusões

Neste Capítulo foi apresentada a formulação matemática do modelo de tráfego desenvolvido no Capítulo 4. A aplicação dos *software Xpress-MP* e *ILOG Cplex* permitiram a obtenção de valores globais ótimos. Os resultados obtidos e as comparações realizadas entre o algoritmo descentralizado e o

¹Em programação linear a função $\text{Min}\{a, b\}$ pode ser escrita como:

$$\begin{aligned} \text{Seja: } & 0 \leq a \leq C, \\ & 0 \leq b \leq C, \quad \text{onde } C \text{ é uma constante} \\ & u \leq a, \quad u = \min\{a, b\} \\ & u \leq b, \quad u \geq b - C(1 - y), \quad y \in \{0, 1\} \\ & u \geq a - Cy \end{aligned}$$

modelo global, permitiram concluir que os resultados com o algoritmo descentralizado aproximam-se em algumas situações de valores do ótimo global, ficando como perspectivas de trabalhos futuros ajustes às heurísticas aplicadas com o objetivo de aproximar os resultados do algoritmo descentralizado ao ótimo global.

Pode-se concluir dos resultados obtidos que o modelo global apresenta como grande limitação o tempo computacional para obtenção de resultados. Um estudo detalhado do tempo computacional consumido nas simulações com o modelo global não foi realizado. No entanto, extensões como a aplicação da técnica de horizonte deslizando e um algoritmo de programação dinâmica poderiam auxiliar para diminuição do tempo computacional consumido.

Capítulo 6

Resultados para malha viária aumentada

6.1 Introdução

No Capítulo 5 foi apresentado o modelo global para uma rede de tráfego e sua descrição em programação matemática. Foram realizadas simulações sobre uma rede de três interseções e seus resultados avaliados. Neste capítulo as simulações serão realizadas sobre uma rede de seis interseções, com duplo sentido de fluxo. Sobre esta rede será aplicado o algoritmo descentralizado de controle em tempo real, o modelo global em programação matemática e o TRANSYT. As comparações entre os resultados serão apresentadas e avaliadas neste capítulo.

6.2 Características da Rede de Testes

Foram realizados testes em uma rede de seis interseções com duplo sentido de fluxo na via arterial mostrada na Figura 6.1. A via arterial tem comprimento total de aproximadamente 1012 m dividida em 4 arcos de aproximadamente 253 m cada. Os arcos secundários são de aproximadamente 350 m e os arcos secundários paralelos à via arterial medem 253 m aproximadamente. A velocidade dos veículos é considerada constante a um valor de 16,7 m/s. As taxas de conversão entre arcos, consideradas constantes e conhecidas. Um exemplo com todas as taxas de conversão e condições iniciais para esta rede de seis interseções, pode ser encontrado no Apêndice A. A distribuição de fluxo entre arcos arteriais e secundários é a mesma utilizada para a rede de três interseções, apresentada na tabela 4.1. Os valores de carregamento testados são dois, alto carregamento (A), ou seja, 90% da capacidade total de chegadas em uma interseção e médio carregamento (B), 60% da capacidade total da interseção. Esta capacidade é considerada 1800 v.p.h. para cada interseção. Como a comparação é realizada com o TRANSYT e com o modelo de tráfego descentralizado com tempos fixos foi escolhida a distribuição de chegadas de veículos constante, pois é o tipo de distribuição utilizada pelo TRANSYT.

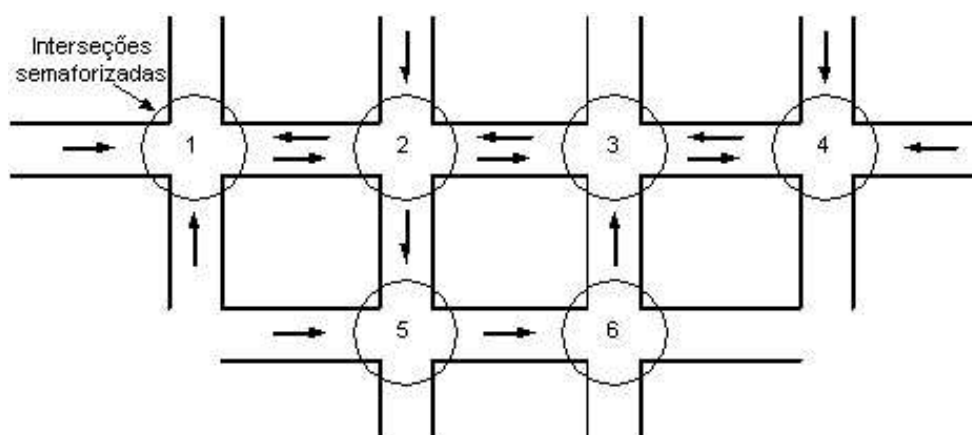


Figura 6.1: Rede de Tráfego.

6.3 Comparações entre Algoritmo Descentralizado e TRANSYT

Inicialmente será apresentada na Tabela 6.1, o valor do atraso obtido diretamente da simulação realizada com o TRANSYT/10 em $pcu-h/h$; na segunda coluna são apresentados os valores do modelo de tráfego descentralizado com tempos fixos (calculados pelo TRANSYT/10 e obtidos na mesma simulação do resultado da primeira coluna desta tabela), em $veíc.-s/s$, e na terceira coluna os valores obtidos diretamente do algoritmo de controle aplicado sobre o modelo de tráfego, em $veíc.-s/s$. Como no Capítulo 4 as comparações foram realizadas com a distribuição de chegadas constante para satisfazer as especificações de entrada do TRANSYT/10. O mínimo tamanho de ciclo aceito pelo *software* TRANSYT/10 é de 30 segundos, que é um valor maior do que a média obtida com o algoritmo descentralizado. Os resultados são de 30 minutos de simulação.

Tabela 6.1: Atraso em $pcu-h/h$ (TRANSYT) e $pcu-s/s$ (descentralizada e fixa).

	TRANSYT	Fixo	Desc.
BA	67,6	39,3	15,7
DA	50,2	28,6	15,3
BM	28,9	21,4	7,4
DM	15,7	11,8	5,6

O resultado apresentado na tabela 6.1 era esperado pois, como no caso de três interseções apresentado no capítulo 4, o algoritmo descentralizado por ser um algoritmo atuado, adapta-se à demanda em tempo real, proporcionando a melhoria apresentada nos resultados. Verifica-se que a diferença entre os valores da primeira e segunda colunas aumentou em relação ao resultado obtido com a rede de 3 interseções. Acredita-se que este comportamento deve-se ao fato de uma rede maior com duplo sentido e um ciclo entre as interseções 2,3,5 e 6, (ver Figura 6.1) causar um maior grau de dificuldade nas predições de fluxo veicular nas vias.

Na tabela 6.2 pode-se verificar o ganho comparativo do algoritmo descentralizado sobre o TRANSYT. Os resultados mostram que o algoritmo descentralizado apresenta ganho elevado sobre o TRANSYT.

Tabela 6.2: Proporção de melhoria do algoritmo descentralizado \times TRANSYT.

Desc./TRANSYT	
BA	330
DA	228
BM	290
DM	180

A proporção de melhoria foi calculada pela expressão:

$$\text{Proporção} = \frac{T - D}{D} 100 \quad (6.1)$$

onde, T é o valor de atraso obtido com o TRANSYT e D o valor de atraso obtido com o algoritmo descentralizado.

Tabela 6.3: Proporção de melhoria do algoritmo descentralizado \times tempo fixo.

Desc./Fixo	
BA	150
DA	87
BM	189
DM	111

Na tabela 6.3 pode-se verificar o ganho comparativo do algoritmo descentralizado sobre o de tempo fixo aplicado ao mesmo modelo. São aplicados os tempos fixos obtidos com o TRANSYT. Os resultados mostram que o algoritmo descentralizado apresenta ganho elevado sobre a estratégia de tempo fixo. A proporção de melhoria foi calculada pela expressão:

$$\text{Proporção} = \frac{D - F}{F} 100 \quad (6.2)$$

onde, D o valor de atraso obtido com o algoritmo descentralizado e F é o valor de atraso obtido com o modelo de tráfego com os tempos fixos do TRANSYT.

Para analisar a formação de filas e a progressão de pelotões será apresentada a Figura 6.2 onde pode-se verificar a adequação do algoritmo descentralizado à demanda verificada e conseqüentemente a obtenção de um melhor desempenho, formando filas de tamanhos menores. As simulações foram realizadas para distribuição de chegada constante, balanceada, alta (BA).

A Figura 6.2 mostra que o algoritmo descentralizado calcula tempos de verde menores do que o TRANSYT, que tem o comprimento de ciclo fixo. Foram testados diferentes comprimentos de ciclo no TRANSYT, no entanto, o melhor resultado é o apresentado. Os dados de entrada aplicados ao TRANSYT e a configuração da rede para o TRANSYT podem ser encontrados no Apêndice C.

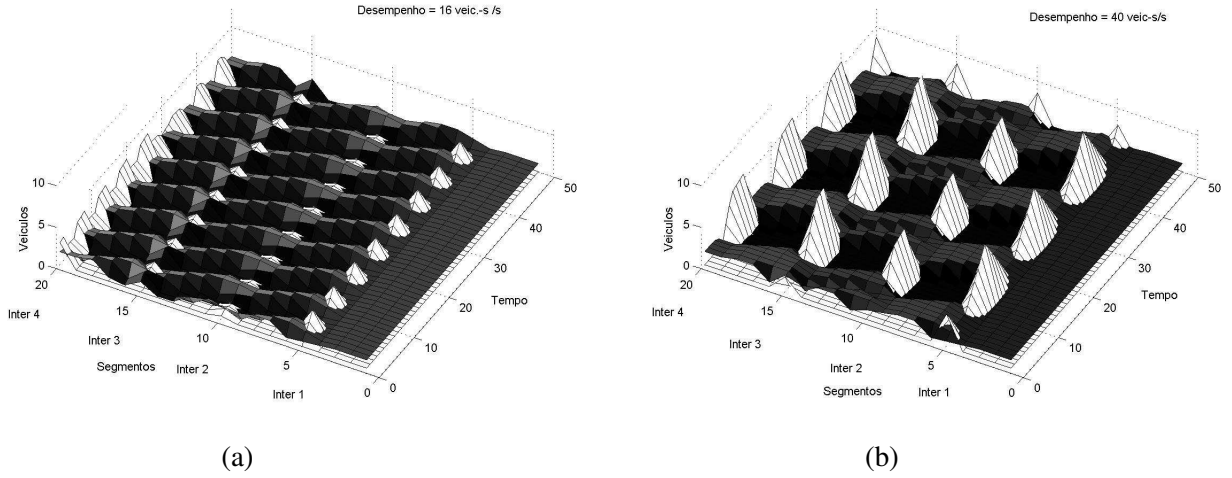


Figura 6.2: Comparação algoritmo descentralizado (a) × TRANSYT (b)

6.4 Comparação do Algoritmo descentralizado com o Modelo Global

A Tabela 6.4 apresenta os valores de atraso obtidos de simulações entre o algoritmo descentralizado e o modelo global. Por limitações de capacidade computacional o modelo global só pode ser simulado para um horizonte de 20 amostras de tempo. Outra limitação foi o tempo de processamento; nenhuma das amostras testadas foi processada até a obtenção do ótimo global. Foi estipulado um limite de 20 horas de processamento, e os resultados para todas as instâncias testadas atingiram valores de melhor desempenho do que para o algoritmo descentralizado.

Tabela 6.4: Atraso (s) - (período de simulação: 20 amostras de tempo) Comparação entre modelos Descentralizado(D) e Global(G).

	Escada		Exponencial		Pulsado		Constante	
	G	D	G	D	G	D	G	D
BA	1213	1365	1148	1422	1198	1408	1129	1314
DA	1093	1227	1317	1444	1130	1331	1307	1342
BM	558	582	565	595	550	582	597	633
DM	460	480	389	489	429	522	413	478

Na Tabela 6.5 pode-se verificar a proporção de melhoria do desempenho encontrado com o modelo global sobre o algoritmo descentralizado. Os valores variam de um mínimo de 2,7% a um máximo de 25,7%. A proporção apresentada na Tabela 6.5 foi calculada como:

$$\text{Proporção} = \frac{D - G}{G} 100 \quad (6.3)$$

Mesmo não tendo sido obtidos valores de ótimo global para as instâncias testadas, os valores de desempenho são melhores do que os valores obtidos com o algoritmo descentralizado. Este resultado era esperado em função das simplificações realizadas através das heurísticas aplicadas ao algoritmo descentralizado. Para a rede testada no Capítulo 4, o algoritmo descentralizado atingiu valores de

Tabela 6.5: Proporção de melhoria do modelo global e algoritmo descentralizado.

	Escada	Exponencial	Pulsada	Constante
BA	12,6	23,9	17,5	16,4
DA	12,2	9,6	17,8	2,7
BM	4,4	5,4	5,8	6,0
DM	4,3	25,7	21,6	15,7

ótimo global para as instâncias de distribuição constante, com carregamento balanceado alto (BA) e para a distribuição pulsada, com carregamento desbalanceado médio (DM). Para a rede de seis interseções este resultado não foi reproduzido, pois a rede possui ciclos, ou seja, chegadas de veículos advindos de descargas de filas que produzem um grau maior de dificuldade ao processo de predição de chegadas.

Através das Figuras 6.3 (a) e (b) pode-se verificar como evolui a busca do ótimo global no processo de otimização da programação matemática aplicada ao modelo global. As Figuras representam dois casos distintos de comportamento da busca. No caso (a) o modelo global encontra um valor sub-ótimo melhor do que o valor obtido pelo algoritmo descentralizado já nos primeiros passos da busca. No caso (b), o valor sub-ótimo encontrado inicialmente é superior ao valor encontrado pelo algoritmo descentralizado, no entanto, após algumas horas de simulação, o valor sub-ótimo torna-se menor do que o obtido pelo algoritmo descentralizado. Os dois casos apresentados são representativos quanto aos casos simulados, pois todos seguem o mesmo padrão de comportamento. Também para todos os casos apresentados na tabela 6.4 a simulação foi realizada até aproximadamente 20 horas, não tendo sido possível provar que o valor obtido é o valor de ótimo global para nenhum dos casos testados.

A Figura 6.4 mostra uma instância de distribuição de chegadas exponencial, carregamento médio, desbalanceado.

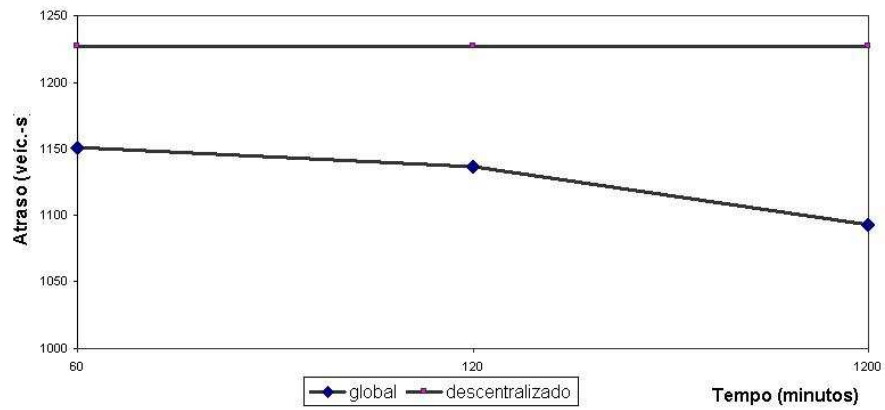
Na Figura 6.4 pode-se verificar o comportamento de descarga de pelotões através da formação de vales. O modelo global apresenta melhor desempenho do que o algoritmo descentralizado. Na Figura (a), pode-se verificar uma maior formação de filas, representada pelos picos brancos.

6.5 Conclusões

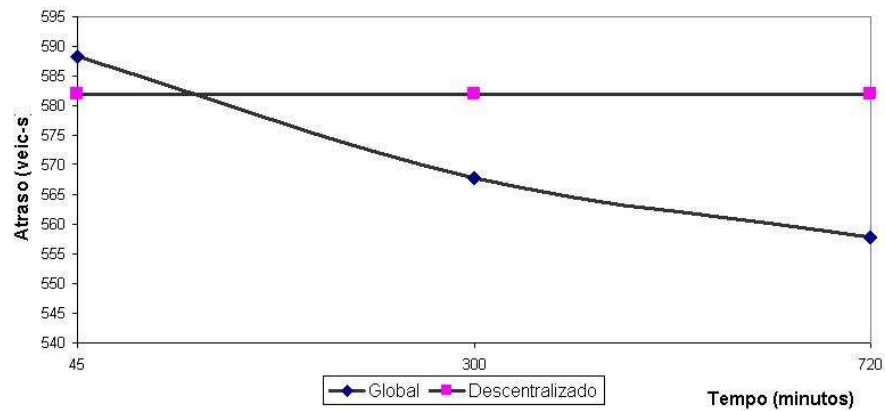
Neste capítulo foi feita uma comparação entre o algoritmo descentralizado, o TRANSYT e o modelo global para uma rede de seis interseções. A comparação com o TRANSYT teve como objetivo proporcionar um índice de melhoria de uma metodologia de controle atuado em tempo real sobre outra de tempo fixo. Já a comparação com o modelo global descrito em programação matemática apresentou o quanto as heurísticas aplicadas ao algoritmo descentralizado podem ser aprimoradas.

Verifica-se também a necessidade de validação do algoritmo descentralizado e do modelo global. Esta validação pode ser realizada através de micro-simulação utilizando-se o simulador SITRA-B+ [7] que tem como objetivo representar, de forma tão precisa quanto possível, o tráfego de veículos em

uma rede urbana considerando, entre outras características, a geometria das interseções, o comportamento típico dos motoristas, e a ocorrência da chegada de ônibus. O simulador proporciona para o usuário uma ferramenta de avaliação do desempenho do sistema viário, que permite a comparação de seus resultados com os de outros tipos de controle de tráfego ou estratégias de controle [7].



(a)



(b)

Figura 6.3: (a) Carregamento desbalanceado, alto, escada, (b) Carregamento balanceado, médio, escada.

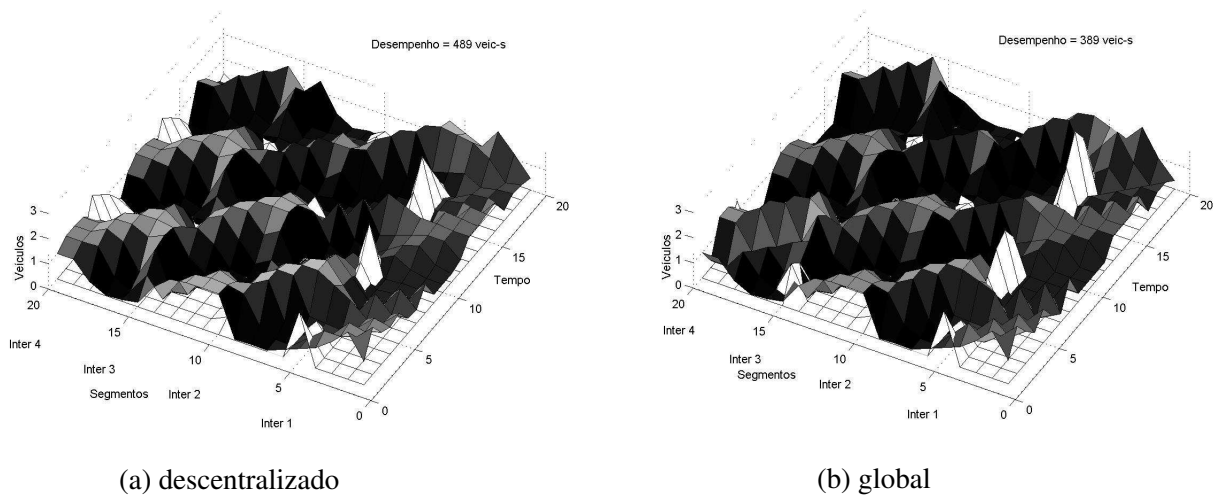


Figura 6.4: Comparação algoritmo descentralizado (a) \times (b) modelo global.

Capítulo 7

Conclusões e Perspectivas

7.1 Introdução

Este trabalho teve como objetivo o desenvolvimento de um algoritmo de otimização do desempenho de uma rede de tráfego urbano. O algoritmo faz o controle dos tempos semafóricos em tempo real em resposta às variações de demanda do sistema. Para aplicação em tempo real foram introduzidas heurísticas ao modelo de tráfego, desenvolvido e aplicado neste trabalho. Heurísticas também foram aplicadas ao algoritmo de controle de forma a proporcionar resultados sub-ótimos para minimização do atraso veicular. Uma destas heurísticas determina a arquitetura descentralizada do algoritmo. Para avaliação destes resultados sub-ótimos obtidos com o algoritmo descentralizado foi desenvolvido um modelo global, onde são consideradas as interferências mútuas entre interseções, descrito por programação matemática e aplicado a *software* comerciais para busca de soluções ótimas globais. O modelo global além de proporcionar uma ferramenta para comparação dos resultados de desempenho, também permitiu avaliação quanto à descarga de veículos e coordenação de pelotões entre interseções.

7.2 Conclusões

O algoritmo descentralizado desenvolvido teve uma avaliação satisfatória, pois nos resultados apresentados foi possível verificar a obtenção de valores sub-ótimos próximos às soluções ótimas globais. Em comparações com a estratégia TRANSYT para as duas redes testadas, também pode ser verificada a melhoria do desempenho.

Os resultados quanto à descarga de filas e coordenação de pelotões teve uma avaliação positiva, pois os resultados demonstraram que o algoritmo descentralizado proporciona a formação de vales caracterizando a descarga de pelotões de forma periódica. Os melhores resultados obtidos foram para a distribuição de chegadas pulsadas, de onde pode-se imaginar que uma regra de limite de velocidade aos usuários, de forma a mantê-los em pelotões, reduziria consideravelmente o atraso veicular. Tal regra precisa ser comprovada como uma perspectiva de trabalhos futuros.

7.3 Perspectivas

Proporcionar uma metodologia para prever bloqueios de vias a montante por existência de fila na via a jusante da interseção. Um caminho estudado na tentativa de solucionar o problema foi a modelagem de filas horizontais, no entanto, não foram obtidos resultados conclusivos, deixando-se esta linha de pesquisa como perspectiva futura.

A coordenação de pelotões e a descarga de filas verificada com o algoritmo descentralizado foram obtidas através de uma metodologia implícita de coordenação. Estudos mostram a existência de metodologia explícitas para tratar este problema. Esta é outra linha de pesquisa a ser considerada no futuro.

As expansões sugeridas ao modelo global descrito em programação matemática também são perspectivas para desenvolvimentos futuros.

Capítulo 8

Publicações

1. S. G. S. Cervantes, W. Kraus Junior. Programação dinâmica aplicada ao controle de tráfego urbano. In: XVI Congresso de Pesquisa e Ensino em Transportes, 2002, Natal, RN. Anais do XVI ANPET.
2. S. G. S. Cervantes, W. Kraus Junior, J. L. Farges. Um algoritmo de busca em profundidade para controle ótimo em tempo real de tráfego urbano. In: XIII Panamerican Congress of Transportation, 2004, Albany, NY. Proceedings of the XIII Panamerican Congress of Transportation, v. 1. p. 12-23.
3. E. Camponogara, S. G. S. Cervantes, W. Kraus Junior. A mathematical programming model for urban traffic control. In: XIII Panamerican Congress of Transportation, 2004, Albany, NY. Proceedings of the XIII Panamerican Conference of Transportation, v. 1. p. 1-12.
4. S. G. S. Cervantes, E. Camponogara, W. Kraus Junior. Programação matemática para controle de tráfego urbano. In: XVIII Congresso de Pesquisa e Ensino em Transportes, 2004, Florianópolis. Anais do XVIII ANPET, v. 1. p. 526-536.


```

# segmento da via l2

param p {(l1,l) in U};      # p(l1,l2) é a proporção de conversão
                             # da l1 para a via l.

set I within {1..L cross 1..L}; # (l1,l2) pertencem a I se as vias
                             # l1 e l2 estiverem em conflito.

set LI within 1..L;

param An {t in 1..T, l in 1..L} >= 0;
      # An(l,t) é o número de veículos que entram na via
      # l até o tempo t

#variáveis #-----

var x {l in 1..L, t in 0..T} >=0;  # número de veículos em fila na
      # via l

var m {l in 1..L, t in 0..(T-1)} integer>=0, <=1;

var a {l in 1..L, j in 1..(N[l]+1), t in 0..T} >= 0;

var y {l in 1..L, t in 0..(T-1)} >= 0;

var z {l in Wy, t in 0..(T-1)} >= 0;

var mq {l in 1..L, t in 0..(T-1)} >= 0;

#-----
# Função Objetivo
#-----

minimize atraso total: 4*sum{l in 1.. L} (0.5*x[l,0] + sum{t in
1..T-1} (0.5*(x[l,t] + x[l,t+1])) + x[l,T]);

#-----
#Restrições
#-----

sujeito a
ini_x {l in 1..L}:
      x[l,0] = x0[l];

sujeito a
ini_a {l in 1..L, j in 1..(N[l]+1)}:

```

```

a[l,j,0] = a0[l,j];

#-----
# x's dinâmicos
#-----

sujeito a
x_dinâmicos_2 {l in 1..L, t in 1..T}:
    x[l,t] >= x[l,t-1] + a[l,1,t-1] - y[l,t-1];

sujeito a
y_dinâmicos_1 {l in 1..L, t in 0..(T-1)}:
    y[l,t] <= x[l,t] + a[l,1,t];

sujeitos a
y_dinâmicos_2 {l in 1..L, t in 0..(T-1)}:
    y[l,t] <= S[l]*m[l,t];

#-----
# a's dinâmicos
#-----

sujeitos a
a_dinâmicos_1 {l in 1..L, j in 1..(N[l]-1), t in 1..T}:
    a[l,j,t] = a[l,j+1,t-1];

sujeitos a
a_dinâmicos_2 {l in Wy, t in 1..T}:
    a[l,N[l],t] = a[l,N[l]+1,t-1] + (1-R[l])*z[l,t-1];

sujeitos a
a_dinâmicos_3 {l in Wy, t in 1..T}:
    a[l,N[l]+1,t] = R[l]*z[l,t-1];

sujeitos a
a_dinâmicos_4 {l in Wn, t in 1..T}:
    a[l,N[l],t] = a[l,N[l]+1,t-1];

sujeitos a
a_dinâmicos_5 {l in Wn, t in 1..T}:
    a[l,N[l]+1,t] = An[t,l];

#-----
# z's dinâmicos
#-----

sujeitos a
z_dinâmicos_2 {l in Wy, t in 0..(T-1)}:
    z[l,t] = sum{(ll,1) in U} p[ll,l]*y[ll,t];

```

```

#-----
# conflito entre indicações verdes
#-----

sujeito a
sinais_verde {l in LI, t in 0..(T-1)}:
    (m[l,t] + sum{(l,ll) in I} m[ll,t]) <= 1;

#-----
# restrições de verde mínimo
#-----

sujeitos a
min_verde_0 {l in 1..L}:
    mq[l,0] = 4;

sujeito a
min_verde_1 {l in 1..L, t in 1..(T-1)}:
    mq[l,t] <= mq[l,t-1] + m[l,t];

sujeito a
min_verde_2 {l in 1..L, t in 1..(T-1)}:
    mq[l,t] >= mq[l,t-1] + m[l,t] - bigZ*(1 - m[l,t]);

sujeito a
min_verde_3 {l in 1..L, t in 1..(T-1)}:
    mq[l,t] <= bigZ*m[l,t];

sujeito a
min_verde_4 {l in 1..L, t in 1..(T-1)}:
    3*m[l,t] >= 3 - mq[l,t-1] - 3*bigZ*(1 - m[l,t-1]);

```

A.1.1 Arquivo de dados - 3 interseções

```
param L := 6; param T := 25;
```

```
param bigZ := 26;
```

```

param:  x0      N      R      S :=
      1  1.5840 3      0.8 2
      2  0.6024 3      0.8 2
      3  2.2896 3      0.8 2
      4  0.9160 5      0.2 2
      5  0.0     5      0.2 2
      6  0.0     5      0.2 2;

```

```

param:  a0 :=
      1 1      0
      1 2      0

```

```

1 3      0.3960
1 4      1.5840

```

```

2 1      1.7820
2 2      1.7820
2 3      1.4656
2 4      0.1600

```

```

3 1      1.7698
3 2      1.3514
3 3      0.1620
3 4      0.1296

```

```

4 1      1.6200
4 2      0.3240
4 3      0
4 4      0
4 5      1.2960
4 6      0.3240

```

```

5 1      1.6200
5 2      0.3240
5 3      0
5 4      0
5 5      1.2960
5 6      0.3240

```

```

6 1      1.6200
6 2      0.3240
6 3      0
6 4      0
6 5      1.2960
6 6      0.3240;

```

```
set Wy := 2 3; set Wn := 1 4 5 6;
```

```
set U := (1,2) (4,2) (2,3) (5,3);
```

```

param: p :=
    1 2      0.9
    4 2      0.1
    2 3      0.9
    5 3      0.1;

```

```
set LI := 1 2 3; set I := (1,4) (2,5) (3,6);
```

```
param An: 1 2 3 4 5 6
```

:=

1	1.9800	0	0	1.6200	1.6200	1.6200
2	1.9800	0	0	1.6200	1.6200	1.6200
3	0	0	0	0	0	0
4	0	0	0	0	0	0
5	0	0	0	0	0	0
6	1.9800	0	0	1.6200	1.6200	1.6200
7	1.9800	0	0	1.6200	1.6200	1.6200
8	1.9800	0	0	1.6200	1.6200	1.6200
9	0	0	0	0	0	0
10	0	0	0	0	0	0
11	0	0	0	0	0	0
12	1.9800	0	0	1.6200	1.6200	1.6200
13	1.9800	0	0	1.6200	1.6200	1.6200
14	1.9800	0	0	1.6200	1.6200	1.6200
15	0	0	0	0	0	0
16	0	0	0	0	0	0
17	0	0	0	0	0	0
18	1.9800	0	0	1.6200	1.6200	1.6200
19	1.9800	0	0	1.6200	1.6200	1.6200
20	1.9800	0	0	1.6200	1.6200	1.6200
21	0	0	0	0	0	0
22	0	0	0	0	0	0
23	0	0	0	0	0	0
24	1.9800	0	0	1.6200	1.6200	1.6200
25	1.9800	0	0	1.6200	1.6200	1.6200;

A.1.2 Arquivo de dados - 6 interseções

```
param L := 16; param T := 20;
```

```
param bigZ := 21;
```

```
param:  x0      N      R      S :=
  1    0.0000  3      0.8  2
  2    0.7960  3      0.8  2
  3    0.0000  3      0.8  2
  4    0.7981  3      0.8  2
  5    0.0000  5      0.2  2
  6    0.0926  3      0.8  2
  7    0.2640  5      0.2  2
  8    0.0000  5      0.2  2
  9    1.0129  5      0.2  2
 10    0.0000  3      0.8  2
 11    0.5940  3      0.8  2
 12    0.9340  5      0.2  2
 13    0.0000  3      0.8  2
 14    0.1531  3      0.8  2
 15    0.0000  3      0.8  2
 16    1.4720  3      0.8  2;
```

```
param: a0 :=
```

```
1 1      0.4640
1 2      0.8000
1 3      0.0660
1 4      0.2640
```

```
2 1      0.0813
2 2      0.3653
2 3      1.6913
2 4      1.0051
```

```
3 1      0.8305
3 2      0.3842
3 3      0.1958
3 4      0.1432
```

```
4 1      0.4147
4 2      1.4715
4 3      0.6915
4 4      0.3041
```

```
5 1      0.0000
5 2      0.0000
5 3      0.0000
5 4      1.2800
5 5      1.4656
5 6      0.2864
```

```
6 1      0.5346
6 2      0.2776
6 3      0.1818
6 4      0.0870
```

```
7 1      0.6020
7 2      0.6700
7 3      0.9340
7 4      0.4640
7 5      0.3300
7 6      0.0660
```

```
8 1      1.3300
8 2      0.8020
8 3      0.9340
8 4      0.2000
8 5      0.0000
8 6      0.0000
```

```
9 1      0.0650
9 2      0.7698
```


9 3	0.5760
9 4	0.5789
9 5	0.2806
9 6	0.0400

10 1	0.0660
10 2	0.2640
10 3	0.2000
10 4	0.8000

11 1	0.2000
11 2	0.9340
11 3	0.7360
11 4	0.8000

12 1	0.4640
12 2	1.4020
12 3	0.8700
12 4	0.3980
12 5	0.6020
12 6	0.1340

13 1	0.3185
13 2	0.1379
13 3	0.1958
13 4	0.1432

14 1	0.4053
14 2	1.6000
14 3	1.4520
14 4	0.6881

15 1	0.1258
15 2	0.1866
15 3	0.1171
15 4	0.0429

16 1	1.1980
16 2	0.6700
16 3	0.7360
16 4	0.8000;

```
set Wy := 2 3 4 5 6 9 13 14 15;  
set Wn := 1 7 8 10 11 12 16;
```

```
set U := (1,2)  
         (2,3) (2,5)  
         (3,4)  
         (5,6)  
         (6,9)  
         (7,2)
```

```

(8,3) (8,5) (8,13)
(9,4) (9,14)
(10,15)
(11,6)
(12,9)
(14,13)
(15,14)
(16,15);

```

```
param: p :=
```

```

1 2 0.9
2 3 0.9
2 5 0.1
3 4 1.0
5 6 0.1
6 9 0.1
7 2 0.1
8 3 0.1
8 5 0.8
8 13 0.1
9 4 0.1
9 14 0.1
10 15 0.1
11 6 0.9
12 9 0.9
14 13 1.0
15 14 0.9
16 15 0.9;

```

```
set LI := 1 2 3 4 5 6 7 8 9 10;
```

```
set I := (1,7) (2,8) (3,9)
```

```
(4,10) (5,11) (6,12) (7,13) (8,14) (9,15) (10,16);
```

```
param An :1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16:=
```

```

1 0.00 0.00 0.00 0.00 0.00 0.00 1.00 0.00 0.00 0.33 0.00 1.33 0.00 0.00 0.00 1.00
2 0.67 0.00 0.00 0.00 0.00 0.00 0.00 0.33 1.00 0.00 1.00 1.00 1.00 0.00 0.00 0.00 0.00
3 0.33 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.33 0.00 0.33 0.67 0.00 0.00 0.00 0.00 0.67
4 0.67 0.00 0.00 0.00 0.00 0.00 0.00 0.67 0.00 0.00 1.33 0.33 0.67 0.00 0.00 0.00 1.00
5 0.67 0.00 0.00 0.00 0.00 0.00 0.00 1.00 0.00 0.00 0.33 0.33 0.67 0.00 0.00 0.00 0.67
6 0.67 0.00 0.00 0.00 0.00 0.00 0.00 1.00 1.33 0.00 0.67 0.67 0.67 0.00 0.00 0.00 0.00
7 0.33 0.00 0.00 0.00 0.00 0.00 0.00 0.33 0.00 0.00 0.00 0.33 0.33 0.00 0.00 0.00 0.67
8 0.33 0.00 0.00 0.00 0.00 0.00 0.00 1.00 0.33 0.00 0.67 0.33 0.33 0.00 0.00 0.00 0.67
9 0.33 0.00 0.00 0.00 0.00 0.00 0.00 0.33 0.00 0.00 1.33 0.67 0.67 0.00 0.00 0.00 1.33
10 1.33 0.00 0.00 0.00 0.00 0.00 0.00 0.33 0.67 0.00 0.00 0.33 0.67 0.00 0.00 0.00 1.00
11 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 1.00 0.67 0.67 0.00 0.00 0.00 1.00
12 0.67 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.33 0.00 0.67 1.67 0.67 0.00 0.00 0.00 1.33

```

13	0.67	0.00	0.00	0.00	0.00	0.00	1.33	0.33	0.00	0.67	0.00	1.00	0.00	0.00	0.00	0.67
14	1.00	0.00	0.00	0.00	0.00	0.00	0.67	0.33	0.00	0.00	0.33	0.33	0.00	0.00	0.00	1.33
15	0.67	0.00	0.00	0.00	0.00	0.00	0.33	0.33	0.00	0.33	0.33	0.00	0.00	0.00	0.00	0.33
16	0.67	0.00	0.00	0.00	0.00	0.00	0.67	0.00	0.00	0.33	1.00	0.33	0.00	0.00	0.00	0.67
17	0.67	0.00	0.00	0.00	0.00	0.00	0.33	0.67	0.00	0.33	0.67	0.00	0.00	0.00	0.00	0.00
18	1.67	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.67	1.33	0.00	0.00	0.00	0.00	0.33
19	0.67	0.00	0.00	0.00	0.00	0.00	0.33	0.33	0.00	0.33	1.00	0.67	0.00	0.00	0.00	1.00
20	1.67	0.00	0.00	0.00	0.00	0.00	0.67	0.67	0.00	1.00	0.33	1.33	0.00	0.00	0.00	0.00;

Apêndice B

Distribuições

Será apresentado neste apêndice as diferentes distribuições de chegadas de veículos em uma rede de tráfego urbano. As distribuições obedecem as quantidades totais especificadas na tabela 4.1 para um carregamento alto, 90% da capacidade total da via, (considerada constante e igual a 1800 v.p.h.) e para um carregamento médio, 60%.

B.1 Distribuição Escada

A distribuição escada representa uma chegada de veículos randômica, onde são atribuídos valores inteiros entre 0 e 2, que representam os valores de chegada de veículos mínima e máxima, por amostra de tempo. Os intervalos de distribuição são definidos para obtenção do valor total de chegadas de veículos por hora. Para um carregamento alto e balanceado a distribuição é apresentada na figura B.1 e o pseudo-código do algoritmo de geração de veículos pode ser descrito da seguinte forma:

Distribuição Escada

```
Faça  $j = 1$  : tempo total de simulação - 1
  Faça  $l = 1$  : número total de vias
     $g =$  valor randômico,  $g \in \{0, 1\}$ 
    Se  $l =$  via 1
      Se  $0 \leq g < 0.15$ 
        chegada = 0
      Se  $0.15 \leq g < 0.41$ 
        chegada = 1
      Se  $0.41 \leq g < 1$ 
        chegada = 2
    Se  $l >$  número total de interseções
      Se  $0 \leq g < 0.64$ 
        chegada = 0
      Se  $0.64 \leq g < 1$ 
        chegada = 1
```

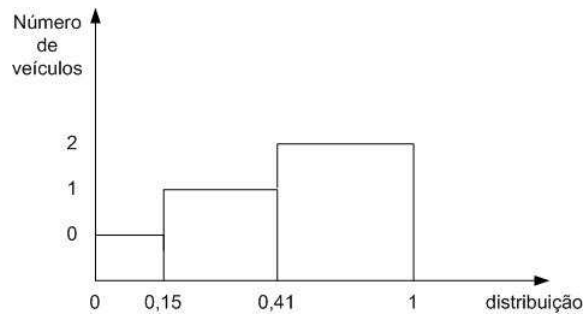


Figura B.1: Distribuição de chegadas para o primeiro arco arterial da rede.

B.2 Distribuição Exponencial

É uma distribuição contínua de probabilidades em que a ocorrência de veículos aumenta ou diminui constantemente. Neste trabalho foi utilizada para representar a chegada de veículos com um aumento constante no tempo.

Os dados de chegada veicular exponencial podem ser obtidos da seguinte forma:

1. são gerados valores randômicos, $r \in \{0, 1\}$;
2. estes valores são aplicados à equação que define uma distribuição exponencial:

$$q = \frac{-1}{\lambda} \log(r) \quad (\text{B.1})$$

onde λ define a quantidade total de veículos que chegam em uma via em um determinado período de tempo, q é o tempo que transcorrido entre duas chegadas;

3. é formado um vetor onde cada célula suporta a quantidade de dois veículos no máximo, acima deste valor, os veículos são transferidos para a célula seguinte.

B.3 Distribuição Constante

A distribuição constante é representada por uma chegada constante de veículos em quantidade suficiente para garantir a demanda definida.

Um exemplo para um carregamento alto e balanceado pode ser descrito da seguinte forma:

Distribuição Constante

Faça $j = 1$: tempo total de simulação - 1

Faça $l = 1$: número total de vias

Se $l = \text{via } 1$

chegada = 0.991

Se $l > \text{número total de interseções}$

chegada = 0.811

B.4 Distribuição Pulsada

A distribuição pulsada representa a alternância entre um valor de chegadas n , onde $0 < n \leq 2$ e *zero*. O valor 2 representa a capacidade máxima da via por amostra de tempo.

Um exemplo para um carregamento alto e balanceado pode ser descrito da seguinte forma:

Distribuição Pulsada

pn = tamanho do período de chegadas nulas

Enquanto $j < 1$: tempo total de simulação -1

Se $j \leq pn$

Faça l = número total de vias

Se l = via 1

chegada = 1.98

Se l > número total de interseções

chegada = 1.62

$j = j + 1$

Se $pn < j \leq pn + 3$

Faça l = número total de vias

Se l = via 1

chegada = 0

Se l > número total de interseções

chegada = 0

$j = j + 1$

Se $j > pn + 3$

$pn = j + 2$

Apêndice C

Configuração do sistema TRANSYT

Serão apresentados os dados de entrada necessários para especificação de simulação no TRANSYT. Estes dados representam um caso de carregamento alto, balanceado para uma distribuição constante. O exemplo aqui apresentado refere-se a rede de seis interseções especificada no capítulo 6.

- tempo de ciclo = 48 segundos;
- número de passo do histogramas de chegadas = 48;
- tempo de simulação = 30 minutos;
- deslocamento do início de verde efetivo = 2;
- deslocamento do final de verde efetivo = 3;
- fator de escala de fluxo = 100;
- fator de escala de tempo = 100;
- otimização de *split* e *offset*;
- fator de dispersão = 35;
- peso % do atraso = 100;
- peso % de parada = 100;
- arco arterial:
 - fluxo total = 891;
 - fluxo uniforme = 0;
 - tempo de cruzamento (fluxo de entrada) = 40;
 - tempo de cruzamento (fluxo de saída) = 20;
- arco secundário:
 - fluxo total = 729;
 - fluxo uniforme = 0;
 - tempo de cruzamento (fluxo de entrada) = 18;
 - tempo de cruzamento (fluxo de saída) = 20;

Referências Bibliográficas

- [1] (1984). *Manual de Semáforos*. Brasília: DENATRAN.
- [2] (1990). *Traffic Network Study Tool - TRANSYT 7F - United States Version*. Mc Trans Center.
- [3] (1997). *Código de Trânsito Brasileiro*. Brasília: Ministério da Justiça.
- [4] (2003). *ILOG CPLEX 9.0: Getting Started*. Mountain View, CA.
- [5] A, J. M. (1963). Setting for fixed-cycle traffic signals. *Operational Research Quarterly*, 14(4):373–386.
- [6] Akcelik, R. (1988). The highway capacity manual delay formula for signalized intersections. *ITE Journal*, 58(3):23–27.
- [7] Bernauer, E., Breheret, L., Algers, S., Boero, M., Taranto, C. D., Dougherty, M., Fox, K., e Gabard, J. (1992). Review of micro-simulation models - appendix d: Analysis of tools - manuel d'utilisation de sitra-b+. Report 042/92, CERT, France.
- [8] Board, T. R. (1998). *Highway Capacity Manual 2000*. National Research Council, Washington D.C., special report 209 Edition.
- [9] Board, T. R. (2000). *Highway Capacity Manual 2000*. National Research Council, Washington D.C., special report 209 Edition.
- [10] Boillot, F. (1994). Evaluation of the real time urban traffic control algorithm CRONOS: first phase. pp 585–590, Tianjin, China. in Proc. 7th IFAC/IFORS Symposium on Transportation.
- [11] Boillot, F., Blossville, J. M., Lesort, J. B., Papageorgiou, M., e Sellam, S. (1992). Optimal signal control of urban traffic networks. pp 75–79, London. in Proc. 6th International Conference on Road Traffic Monitoring and Control, IEE.
- [12] Brito, R. M. (2000). Desenvolvimento de um simulador para análise e projeto de sistemas de controle de tráfego em malha fechada. Dissertação de Mestrado, Universidade Federal de Santa Catarina - UFSC.
- [13] Carlson, R. C. (2004). Implementação de algoritmo para controle em tempo real de tráfego urbano. Projeto de Fim de Curso.
- [14] Coeymans-Avaria, J. e et al (1995). SCOOT in santiago. volume 394, pp 183–195. 23rd European Transport Forum.
- [15] Cormen, T. H., Rivest, R. L., Stein, C., e Leiserson, C. E. (2002). *Algoritmos: Teoria e Prática*. Edição Americana.
- [16] Crabtree, M. R., Vincent, R. A., e Harrison, S. (1996). *TRANSYT/10 User Guide*. Transport Research Foundation Group of Companies, Crowthorne, Berkshire, application /guide 28 Edition.
- [17] D'Azzo, J. J. (1975). *Linear Control System Analysis and Design*. McGraw-Hill, Inc.
- [18] Dell'Olmo, P. e Mirchandani, P. B. (1995). Realband: An approach for real-time coordination of traffic flows on a network. *Transportation Research Record*, 1494:106–116.

- [19] Diakaki, C. (1999). *Integrated Control of Traffic Flow Corridor Networks*. Phd Thesis, Technical University of Crete, Department of Production Engineering and Management.
- [20] Diakaki, C., Papageorgiou, M., e Aboudolas, K. (2002). A multivariable regulator approach to traffic-responsive network-wide signal control. *Control Engineering Practice*, 10:183–195.
- [21] Farges, J. L., Kamdem, I., e Lesort, J. B. (1991). Realization and test of a prototype for real time urban traffic control. Tech report, Drive Project V1022.
- [22] Farges, J. L., Khoudour, L., e Lesort, J. B. (1990). Prodyn: on site evaluation. pp 62–66, Washington, DC. in Proc. Road Traffic Control, 3th International Conference.
- [23] Farines, J., Fraga, J. S., e Oliveira, R. S. (2000). *Sistemas de Tempo Real*. IME-USP, São Paulo-SP.
- [24] Fourer, R., Gay, D. M., e Kernighan, B. W. (2002). *AMPL: A Modeling Language for Mathematical Programming*.
- [25] Garbacz, R. E. (2002). Adaptive signal control: what to expect. Washington, DC. 12th Annual Meeting of the Intelligent Transportation Society of America.
- [26] Gartner, N. H. (1982). Development and testing of a demand-responsive strategy for traffic signal control. pp 578–583.
- [27] Gartner, N. H. (1983). OPAC a demand responsive strategy for traffic signal control. *Transportation Research Record*, 906:75–81.
- [28] Gartner, N. H. (2001). Optimized policies for adaptive control. Washington D.C. Workshop on Adaptive Traffic Signal Control Systems Transportation Research Board.
- [29] Gartner, N. H., Pooran, F. J., e Andrews, C. M. (2001). Implementation of the opac adaptive control strategy in a traffic signal network. *IEEE Intelligent Transportation Systems Conference Proceedings*, pp 195–200.
- [30] Gartner, N. H., Tarnoff, P. J., e Andrews, C. M. (1991). Evaluation of optimized policies for adaptive control strategy. *Transportation Research Record*, 1324:105–114.
- [31] Gazis, D. C. e Potts, R. B. (1963). The oversaturated intersection. pp 221–237, London, UK. Proceedings of the 2nd International Symposium on Traffic Theory.
- [32] Gerlough, D. L. e Huber, M. J. (1975). Traffic flow theory. Tech Report Special Report 165, Transportation Research Board, Washington, D. C.
- [33] Guéret, C., Prins, C., e Sevaux, M. (2002). *Applications of optimization with Xpress-MP*. Englewood Cliffs, NJ.
- [34] Hounsell, N. B. e McDonald, M. (2001). Urban network traffic control. *Institution of Mechanical Engineers*, 215:325–334.
- [35] Hunt, P. B., Robertson, D. I., Bretherton, R. D., e Winton, R. I. (1981). SCOOT - a traffic responsive method of coordinating signals. Report LR 1014, Transport and Road Research Laboratory, Crowthorne, England.
- [36] Hunt, P. B., Robertson, D. L., e Bretherton, R. D. (1982). The scoot on-line traffic signal optimization technique. *Traffic Eng. Control*, 23:190–192.
- [37] Kang, Y. (2000). *Delay, stop and queue estimation for uniform and random traffic arrivals at fixed-time signalized intersections*. Phd Thesis, Faculty of the Virginia Polytechnic Institute and State University, Blacksburg, Virginia.
- [38] Kuester, J. J. e Mize, J. (1973). *Optimization Techniques with Fortran*. New York: McGraw-Hill.
- [39] Lee, S. H. (2005). *Introdução ao Projeto Geométrico de Rodovias*. Editora da UFSC, Florianópolis.
- [40] Little, J. D. C. (1966). The synchronization of traffic signals by mixed-integer linear programming. *Operations Research*, 14:568–594.

- [41] Little, J. D. C., Kelson, M. D., e Gartner, N. H. (1981). MAXBAND: A program for setting signals on arteries and triangular networks. *Transportation Research Record*, 795:40–46.
- [42] Lowrie, P. (1982). The sydney co-ordinated adaptive traffic system - principles, methodology, algorithms. volume 82, pp 67–70. IEE International Conference Road Traffic Signalling.
- [43] Lowrie, P. (1990). SCATS: A traffic responsive method for controlling urban traffic. Tech report, Roads and Traffic Authority, NSW, Australia.
- [44] Ming, S. H. (1997). Nt 201 - uma breve descrição do sistema scoot. Notas Técnicas NT 201, Companhia de Engenharia de Tráfego - CET.
- [45] Mirchandani, P. e Head, L. (2001). Rhodes: A real-time traffic signal control system: architecture, algorithms, and analysis. *Transportation Research – Part C*, 8:105–114.
- [46] Morari, M. e Lee, J. H. (1999). Model predictive control: Past, present and future. *Computers and Chemical Engineering*, 23:667–782.
- [47] Morton, T. e Pentico, D. (1993). *Heuristic Scheduling Systems*. John Wiley and Sons, New York, NY.
- [48] Newell, G. F. (1996). The rolling horizon scheme of traffic signal control. Tech report, Institute of Transportation Studies University of California at Berkeley.
- [49] Papageorgiou, M., Diakaki, C., Dinopoulou, V., Kotsialos, A., e Wang, Y. (2003). Review of road traffic control strategies. *Proceedings of the IEEE*, 91(12):2043–2067.
- [50] Porche, I. e Lafortune, S. (1997). Dynamic traffic control: Decentralized and coordinated methods. IEEE Conference on Intelligent Transportation System.
- [51] Porche, I. e Lafortune, S. (2000). Adaptive look-ahead optimization of traffic signals.
- [52] Porche, I., Sampath, M., Sengupta, R., Chen, Y. L., e Lafortune, S. (1996). A decentralized scheme for real-time optimization of traffic signals. pp 582–589. IEEE International Conference on Control Applications.
- [53] R., W., R., B., R., G., e J., H. (1985). *Traffic Control Systems Handbook*. U.S. Department of Transportation, Washington D.C., special report - fhwa-ip-85-11 Edition.
- [54] Rawlings, J. B. (1999). Tutorial: Model predictive control technology. Em *Proceedings of the American Control Conference*, pp 662–676, San Diego, CA.
- [55] Robertson, D. I. (1968). TRANSYT: A traffic network study tool. Tech report, Road Research Laboratory, Crowthorne, England.
- [56] Robertson, D. I. e Bretherton, R. D. (1991). Optimization networks of traffic signals in real-time: the SCOOT method. *IEEE Transactions on Vehicular Technology*, 40:11–15.
- [57] Robertson, D. I. e Hunt, P. (82). A method of estimating the benefits of coordinating signals by transyt and scoot. *Traffic Engineering and Control*, 23(11):527–531.
- [58] Roupail, N., Tarko, A., e Li, J. (1992). Traffic flow theory and characteristics. Tech report.
- [59] Webster, F. V. (1969). Traffic signal settings. Tech Report 39, Department of Scientific and Industrial Research.
- [60] Williams, H. P. (1999). *Model Building in Mathematical Programming*. John Wiley.
- [61] Wilshire, R., Black, R., Grochowski, R., e Hogonbothan, J. (1985). *Traffic Control System Handbook*. PAWA-Winkelmann and Associates, Inc, 12660 Coit Road, Dallas, TX 75251, institute of transportation engineers Edition.
- [62] Wolsey, L. A. (1998). *Integer Programming*. John Wiley & Sons.